

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS MATEMATICAS
Departamento de Análisis Matemático



TESIS DOCTORAL

Sobre ciertos procesos iterativos isométricos-y protectivos

MEMORIA PARA OPTAR AL GRADO DE DOCTOR
PRESENTADA POR

Manuel Moran Cabre

Madrid, 2015

IT
UCM
1988

T
517.938
Mor

UNIVERSIDAD COMPLUTENSE DE MADRID

Facultad de Ciencias Matemáticas

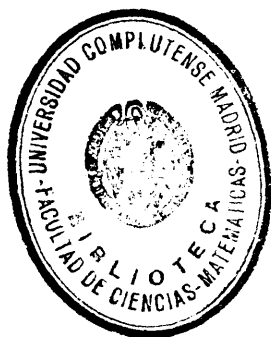
Departamento de Análisis Matemático

BIBLIOTECA UCM



5304845627

**SOBRE CIERTOS PROCESOS
ITERATIVOS ISOMETRICOS Y PROTECTIVOS**



R.35.911

Manuel Morán Cabré

Madrid, 1988

Colección Tesis Doctorales. N.º 486/88

© Manuel Morán Cabré

Edita e imprime la Editorial de la Universidad
Complutense de Madrid. Servicio de Reprografía
Noviciado, 3 - 28015 Madrid
Madrid, 1988
Ricoh 3700
Depósito Legal: M-38859-1988

NC: X-53-017430-7

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE MATEMATICAS

DEPARTAMENTO DE ANALISIS MATEMATICO

"SOBRE CIERTOS PROCESOS ITERATIVOS ISOMETRICOS Y PROTECTIVOS"

TESIS DOCTORAL DE MANUEL MORAN

DIRECTOR DE TESIS : MIGUEL DE GUZMAN

INDICE DE CONTENIDOS

Introducción pag. 1

Capítulo 1 pag. 10

Capítulo 2 pag. 21

Capítulo 3 pag. 53

Anexo de gráficos al capítulo 3 pag. 73

Capítulo 4 pag. 89

Capítulo 5 pag. 111

Anexo de gráficos al capítulo 5 pag. 141

Bibliografía pag. 168

INTRODUCCION

La presente memoria agrupa varios trabajos de investigación que he realizado, bajo la dirección de Miguel de Guzmán, en el departamento de Análisis Matemático de la Universidad Complutense.

* * *

La memoria consta de tres bloques bien diferenciados. El primer bloque comprende el capítulo primero de la tesis, y está dedicado a la confirmación de una conjetura geométrica planteada en 1982 por I.J. Schoenberg (ver ref. [S1]). En rasgos generales, el problema es determinar si cierta cadena de pentágonos planos convexos encajados sucesivamente y cuya construcción está enteramente determinada por el primero de ellos, tiene intersección reducida a un punto. Lo notable de este problema, es que admite una solución natural si se plantea en el plano proyectivo, ya que entonces puede construirse una colineación asociada al pentágono inicial, cuyas iteraciones aplican los vértices de cada pentágono de la cadena en los del siguiente. El estudio de las propiedades iterativas de la colineación, basado en las formas canónicas de Jordan compatibles con las condiciones geométricas del problema, conduce a la contestación afirmativa a la conjetura de Schoenberg, obteniéndose también condiciones para la construcción con regla y compás del punto definido por la cadena de pentágonos. Hemos encontrado un interesante desarrollo de esta idea, que no incluimos en el presente trabajo, consistente en el estudio de las iteraciones de una aplicación lineal proyectiva en el espacio proyectivo complejo n -dimensional.

* * *

El segundo bloque de la tesis, consta de los capítulos II, III y IV. Está dedicado al estudio de ciertos sistemas dinámicos no continuos, que hemos denominado isometrías locales biyectivas — utilizaremos la abreviatura i.l.b. —. Una aproximación intuitiva a dichos sistemas es la siguiente: Supongamos que tenemos un rompecabezas tal que con sus piezas podemos formar la misma figura C de dos maneras diferentes. El cambio de posición de los puntos de C entre la primera y la segunda forma de obtener C con las piezas del rompecabezas, define una transformación $T : C \longrightarrow C$ que es precisamente una isometría local en el sentido de que todo punto de C está incluido en una pieza del rompecabezas que es transportada por T de forma isométrica.

Un caso particular de este tipo de transformaciones son los hoy llamados intercambios de intervalos, cuyo primer antecedente del que tengamos constancia es un trabajo de Oseledecz en 1966 (ver ref. [S2]). Los intercambios de intervalos son isometrías locales en el intervalo $[0,1]$, cuyas "piezas" son intervalos de la forma $[a,b]$ que resultan trasladadas por el intercambio de intervalos, sin permitirse el cambio de orden dentro de cada pieza. Los intercambios de intervalos, han jugado un importante papel en el análisis de ciertos sistemas dinámicos. Ya. G. Sinai (ver ref.

[S2]), que denomina permutaciones a estas transformaciones, explica así su importancia:

« Muchos espacios medibles naturales, son isomorfos como espacios medibles al intervalo $[0,1]$ con la medida Lebesgue usual. Las transformaciones del intervalo que preservan la medida, vienen dadas por ciertas funciones medibles $f: [0,1] \longrightarrow [0,1]$. Tales funciones pueden ser aproximadas por funciones lineales a trozos, que conducen a transformaciones del tipo de las permutaciones (aquí una inclinación -1 corresponde a una traslación y reflexión por simetría del intervalo). »

« De esta forma la permutaciones pueden ser usadas para aproximar transformaciones arbitrarias que preservan la medida en el intervalo. Además, las mismas permutaciones se presentan en ciertos ejemplos interesantes de sistemas dinámicos. »

En 1975, M. Keane encontró condiciones suficientes para la minimalidad de los intercambios de intervalos (ver ref. [K1]), y en 1978, junto a C. Boldrighini y F. Marchetti (ver ref. [K2]), la posibilidad de usar los intercambios de intervalos en el estudio de billares poligonales de ángulos racionales.

Nuevas aplicaciones de los intercambios de intervalos han sido halladas en el estudio de flujos de superficies orientables que preservan el área (ver ref. [M1]). Otra dirección en la que se ha extendido el estudio de los intercambios de intervalos es la determinación de las propiedades de ergodicidad de los mismos, obteniéndose que, en cierto espacio de medida construido para los intercambios de intervalos, y bajo cierta restricción para la permutación de los intervalos, casi todos ellos son únicamente ergódicos. Remitimos para estos resultados a las siguientes referencias bibliográficas: [S2], [K1], [K3], [K4], [K5], [M2], [V1].

Hasta donde sabemos, los intercambios de intervalos son el único modelo de isometría local que haya sido estudiado. Al parecer no existen estudios de otros tipos de isometrías locales en la recta ni en otros espacios más generales.

En el capítulo II se aborda una versión unidimensional de las i.l.b. Se realiza el estudio tomando por dominio de las mismas un subconjunto S de R cualquiera con las únicas condiciones de que sea localmente conexo y acotado, con posible existencia de puntos aislados, y pudiéndose invertir el orden de las piezas. Se supone un número finito de puntos de discontinuidad, y que cada punto no aislado se transporta por la transformación isométricamente junto con todo un intervalo de diámetro positivo. Los intercambios de intervalos, son, desde luego, un caso particular comprendido en este esquema general. Con las hipótesis señaladas, se obtiene la preservación de la medida por estas transformaciones, y se demuestra que, bajo una hipótesis adicional de regularidad, el dominio S puede ser descompuesto en una parte que agrupa todos los puntos periódicos, cuya estructura es completamente analizada, y otra parte que no tiene ningún punto periódico, reduciéndose entonces el

estudio de las r.l.b. al caso de las acíclicas. A continuación se analizan las de este tipo, probándose que en su dominio no puede haber sino un número finito de componentes conexas de diámetro positivo. Se consigue probar que el conjunto de puntos aislados se aplica en sí mismo, y de esta forma se reduce el estudio de las r.l.b. al caso en que el dominio consta de un número finito de componentes conexas de diámetro positivo. Este tipo de transformaciones reciben el nombre de reordenaciones por intervalos — en abreviatura, r.i.—.

Para las r.i. acíclicas se hace un estudio de las propiedades de las órbitas. Se prueba el teorema de recurrencia, versión fuerte (para todo punto) del teorema de recurrencia de Poincaré, bajo cierta hipótesis de regularidad más débil que el criterio suficiente para la condición de minimalidad usado por Keane (ver ref. [K1]) para obtener la minimalidad de los intercambios de intervalos, y se caracterizan, bajo la misma hipótesis, las adherencias de las órbitas de los puntos — que coinciden con los ω -conjuntos — como uniones finitas de intervalos de diámetro positivo (teorema de estructura). Con una hipótesis homologable a la utilizada por Keane se consigue probar que la adherencia de la órbita de un punto coincide con la unión de todas las componentes conexas visitadas por el punto. Resulta así que estas transformaciones pueden descomponerse en sus restricciones a un número finito de partes del dominio en las que se comportan como r.i. minimales (teorema de minimalidad).

Finalmente en el último apartado del capítulo, se extienden estos resultados a un tipo de transformaciones, que son las que se suelen presentar en la práctica, en las que sólo se exige inyectividad en el complementario de un número finito de puntos, para los que tampoco se exige continuidad lateral. Se prueba que la dinámica de estas transformaciones está subordinada a la de cierta r.i., y se concluye el capítulo obteniéndose versiones del teorema de estructura y recurrencia para el caso de r.i. acíclicas en las que no se hace ninguna hipótesis de regularidad.

El fruto de este estudio, es la obtención de un aparato especialmente útil para su aplicación a ciertos sistemas dinámicos. En el capítulo III, se muestra como son las reordenaciones por intervalos las transformaciones que se presentan de forma natural en el problema de los billares poligonales de ángulos racionales. Los intercambios de intervalos no son aplicables directamente, a causa de que no puede garantizarse la preservación del orden a trozos. También parecen las r.i. el aparato adecuado para el estudio de los citados billares, en el caso de que existan obstáculos poligonales. Respecto a la aplicación planteada por Ya. G. Sinaï, queda patente la necesidad de contemplar permutaciones en las que los intervalos se transformen invirtiendo su orden (llamamos en nuestro estudio a este tipo de r.i. no orientables).

* * *

El capítulo III desarrolla un campo de aplicaciones de las transformaciones estudiadas en el capítulo II : Los billares poligonales racionales. Este problema dinámico ha sido bastante estudiado. Las aportaciones de nuestro trabajo, detalladas a continuación, consisten en la solución de ciertos casos singulares que quedaban por resolver, y que completan el estudio del problema de estructura de las órbitas. También aportamos ciertos resultados relativos a la caracterización de órbitas periódicas.

La dinámica de los billares poligonales racionales puede estudiarse a través de cierta reordenación por intervalos. En 1986 (ver ref [K5]) Steven Kerckhoff, Howard Masur y John Smillie establecieron la ergodicidad del sistema salvo para un conjunto de orientaciones iniciales de medida nula. Utilizando el aparato de intercambio de intervalos, M. Keane, Carlo Boldrighini y Federico Marchetti habían conseguido en 1978 (ver ref [K2]) establecer que para las orientaciones iniciales genéricas, todas las órbitas son densas en la mesa de billar, y caracterizaron la familia de orientaciones singulares como una familia numerable, en el caso de billares poligonales racionales de borde conexo. En el apartado 3.6 mostramos, siguiendo el método de Keane, que no obstante la validez de estos resultados, para su correcta obtención se precisa la utilización del aparato de reordenaciones por intervalos no orientables (ver 3.6.2).

Por otra parte, la familia de orientaciones singulares, aunque numerable, es densa en $[0, 2\pi)$, como probamos en 3.6.4 y el trabajo de Keane dejaba abierta la cuestión de la caracterización de la estructura de las órbitas en caso de las orientaciones singulares. De hecho, en 1984 (ver ref. [G1]), Gal'perin ha encontrado ejemplos concretos de órbitas no densas y no periódicas en billares poligonales racionales. En el apartado 3.6 caracterizamos las órbitas correspondientes a orientaciones singulares.

El caso de billares racionales con borde no conexo — billares con obstáculos —, no abordable mediante intercambios de intervalos, también es tratado en el apartado 3.6 por el método de Keane obteniendo, gracias a lo probado sobre isometrías locales en el anterior capítulo, la posibilidad de extender a este tipo de billares los resultados obtenidos en el trabajo de Keane para los billares racionales con borde conexo: Misma caracterización de orientaciones singulares y densidad de la órbitas en la mesa, en el caso de orientaciones genéricas. El resultado general para cualquier órbita en un billar poligonal racional de borde conexo o no, es la densidad de la órbita en una región de la mesa compuesta por un número finito de bandas, que son todas de anchura positiva, o todas reducidas a segmentos, en caso de órbitas periódicas. Hasta donde sabemos, para este tipo de billares el único resultado conocido era el mencionado de Steven Kerckhoff, que significa la ergodicidad, y consecuentemente, la densidad de las órbitas, salvo para un conjunto de orientaciones singulares de medida nula, pero dejaba sin abordar la caracterización de estas últimas, y la estructura de las órbitas que a ellas corresponden.

En otro orden de problemas relacionados con billares poligonales racionales tratamos también en este capítulo algunas cuestiones relativas a las órbitas periódicas. En los billares poligonales racionales, juega un destacado papel el grupo R generado por las simetrías cuyos ejes son los lados del polígono. En el apartado 3.5 establecemos una condición necesaria para que una órbita sea periódica, mostrándose la conexión entre órbitas periódicas y el subgrupo T de traslaciones del grupo R . Formulamos una conjetura sobre la condición suficiente para la periodicidad de las órbitas, que nos ha permitido encontrar órbitas periódicas de hasta varios miles de rebotes mediante computador. En el anexo a este capítulo se muestran algunos ejemplos. En el apartado 3.4 se prueba que el subgrupo T de traslaciones está contenido en cierto grupo abeliano libre generado por una colección de vectores libres (consideramos aquí solamente el caso de billares poligonales racionales de borde conexo), y se estudian varios tipos de mesas en que este grupo es especialmente sencillo, isomorfo concretamente al grupo abeliano libre generado por las raíces p -ésimas de la unidad para valores de p que dependen de la mesa.

Finalmente, en el apartado 3.3 se muestra otro bonito enfoque del problema de los billares poligonales racionales debido a A.N. Zemlyakov y A.B. Katov, que pone de relieve la relación existente entre este problema y el estudio de flujos geodésicos en superficies.

* * *

El bloque destinado a isometrías locales, finaliza con el capítulo IV que está destinado al estudio de ciertas versiones de i.l.b. en el plano. Se introduce una definición de i.l.b. en R^2 basada en el ejemplo natural que nos brindan los billares poliédricos racionales. Para este tipo de transformaciones, no tenemos constancia de la existencia de ningún estudio hasta la fecha.

Se comienza tratando dos casos particulares de i.l.b. en R^2 que pueden ser reducidos a flujos geodésicos en un toro. Se ofrecen algunos resultados generales para las i.l.b. en R^2 , deducibles de la definición. Concretamente se obtiene la recurrencia del sistema, aplicando el teorema de Poincaré, y se prueba que si no existen puntos periódicos, las órbitas de los puntos son densas en círculos, o bien se aproximan tanto como se quiera a los puntos de discontinuidad de la transformación. Un papel importante en este estudio juega el grupo de transformaciones generado por las isometrías que coinciden con las restricciones de la i.l.b. a las componentes isométricas (las "piezas" del rompecabezas). Si en dicho grupo no hay rotaciones irracionales, entonces si la i.l.b. es acíclica, todas las trayectorias pasan arbitrariamente cerca de los puntos de discontinuidad de la i.l.b. Este resultado se aplica a los billares poliédricos racionales, probándose que, salvo para un conjunto numerable de orientaciones, todas las trayectorias se aproximan arbitrariamente a las aristas del poliedro.

No hemos conseguido, sin embargo, obtener resultados homologables al caso unidimensional en lo relativo al problema de la minimalidad

del sistema. De hecho, el flujo racional en un toro, permite construir un ejemplo de i.l.b. en que las órbitas no son ni periódicas ni densas en una región de \mathbb{R}^2 , que sería el resultado homólogo al teorema de estructura probado para el caso unidimensional.

El estudio prosigue centrándose en ciertos tipos sencillos de i.l.b. en \mathbb{R}^2 , particularmente productos cartesianos de reordenaciones por intervalos, y productos cartesianos mixtos. Hemos conseguido clasificar completamente un tipo particular de estos últimos, que constituyen un sistema dinámico equivalente al de los automorfismos del toro que se obtienen al partir el toro en un número finito de rodajas por planos perpendiculares a su eje, y reconstruirlo después situando arbitrariamente estas piezas. También se analiza el caso de producto cartesiano de r.i. en que alguno de los factores son periódicos.

El estudio de i.l.b. en \mathbb{R}^2 concluye buscando condiciones para la minimalidad (más exactamente la casi minimalidad) del producto cartesiano de reordenaciones por intervalos. Encontramos una condición intermedia entre la propiedad mezcladora y ergódica, que llamamos semimezcladora. Se analiza el caso particular $f \times f$, donde f es una reordenación por intervalos minimal, que presenta el interés de que si $f \times f$ no es casi minimal, entonces f no puede ser semimezcladora ni por tanto mezcladora. Se encuentra para este caso un conjunto denso en el dominio, pero de medida nula, tal que los w -conjuntos de sus puntos no cubren todo el dominio (de hecho tienen medida nula), y se utiliza esta idea para la obtención por computadora de gráficos de la r.i. f^k , siendo k un número natural elevado (ver anexo al capítulo). El estudio de la i.l.b. de tipo $f \times f$ presenta el interés de tener consecuencias en el estudio de billares poligonales racionales con dos bolas, aunque estas no son desarrolladas en nuestro trabajo.

El capítulo IV finaliza con una indicación del aparato matemático necesario a desarrollar para el estudio de billares poligonales de ángulos no racionales. Se prueba, usando una idea que Birkhoff utilizó para billares de borde diferenciable (ver ref. [B1]), que puede construirse para billares de este tipo un espacio de fases en que el problema se reduce a un automorfismo diferenciable por piezas, que aunque no preserva las distancias si lo hace con el área, colocándose así el problema de este tipo de billares en una perspectiva de desarrollo de técnicas similares a las de las isometrías locales. Con esto finaliza el capítulo, y el bloque de la tesis dedicado a isometrías locales.

* * *

El tercer bloque de este trabajo, comprende el capítulo V, dedicado al estudio de cierto tipo de algoritmos que permiten la construcción de conjuntos fractales. En una i.l.b., las órbitas de los elementos tienen la propiedad de que los entornos de dos puntos arbitrarios de la órbita (suponiendo que en la misma no haya puntos de discontinuidad), son idénticos a pequeña escala. Tratando de estudiar las posibles propiedades derivadas de este hecho,

más concretamente, buscando otros conjuntos con esta propiedad, obtuvimos los primeros ejemplos de conjuntos asociados a una serie absolutamente convergente, que luego dieron lugar al desarrollo de este capítulo de la tesis.

Probamos que toda serie de vectores de \mathbb{R}^n normalmente convergente tiene asociado un cierto conjunto compacto, que es el conjunto de todas las posibles sumas de términos de la serie. Si la serie tiene infinitos términos no nulos, su conjunto asociado tiene además potencia del continuo, y se trata de un conjunto perfecto. Con cierta hipótesis adicional sobre los restos de la serie de normas, se consigue probar que la medida de Lebesgue del conjunto asociado, es nula. A las series cuyos conjuntos asociados tienen potencia del continuo y medida Lebesgue n -dimensional nula (siendo n la dimensión del espacio vectorial generado por los vectores de la serie), las denominamos series fractales, y como se prueba con varios ejemplos a lo largo del capítulo, de sus conjuntos asociados, cabe esperar comportamiento fractal. Se obtiene una acotación de la dimensión de Hausdorff del conjunto asociado a una serie, que depende de cierta función asociada a las repeticiones de los términos de la serie y de los restos de la serie de normas.

Se analizan los casos particulares de las series geométricas reales y complejas, probándose en estos casos, que acotando oportunamente la norma de la razón, los fractales que se obtienen son autosemejantes, y se obtiene su dimensión de Hausdorff que resulta coincidir con la cota superior hallada anteriormente para el caso general, lo que prueba la economía de los recubrimientos utilizados en la obtención de esta cota. En el caso real, se prueba que la acotación de la razón suficiente para que la medida Lebesgue sea nula, es también necesaria, quedando este problema abierto para el caso complejo. Se muestra la utilidad de las series fractales para la obtención de conjuntos fractales formados por números con expresiones decimales especiales, mostrándose la existencia de una interesante conexión entre la dimensión de Hausdorff y el número de sumas diferentes que pueden ser obtenidas con una colección de números.

A continuación se muestra un procedimiento natural muy amplio para la obtención de series fractales. Se prueba concretamente que para cada función analítica real o compleja en un abierto D , cada punto $x_0 \in D$ tiene un entorno tal que a cada punto del entorno (a excepción del propio x_0) corresponde una serie fractal, con un conjunto fractal asociado que puede variar notablemente para los distintos puntos de un mismo entorno, como se pone de relieve en el anexo de gráficos de este capítulo. Se obtienen las dimensiones de Hausdorff de los fractales asociados a toda la familia de funciones reales $1/(1-x)^p$, $p \in \mathbb{N}$, quedando abierta la determinación exacta de las medidas de Hausdorff de los fractales en sus correspondientes dimensiones (salvo en el caso $p = 1$, en que dichas medidas sí se determinan). Parece probable que la habitual de Hausdorff no sea la más adecuada para este tipo de fractales, y que sea preciso utilizar medidas de Hausdorff generalizadas, como las descritas en 5.1.3. En general presenta mucho interés el estudio de dimensiones y medidas de los fractales que se obtienen mediante

diversos tipos de funciones analíticas, ya que estos prometen una gama muy variada para cuya adecuada clasificación será preciso disponer de un sistema muy fino de medidas.

El capítulo se cierra describiendo otros tipos de algoritmos que, con espíritu similar al de las series fractales, sirven para la obtención de conjuntos fractales. Se hace un estudio particular para el caso de los productos infinitos, obteniéndose condiciones suficientes para que el conjunto asociado al producto infinito sea compacto, perfecto y de medida Lebesgue nula. También se describen otros tipos de algoritmos que sirven para la obtención de fractales basados en series, productos infinitos, fracciones continuas ..., algunos de los cuales permiten la codificación sencilla de fractales conocidos, como el conjunto de Besicovitch, conjuntos autosemejantes, etc. Es notable observar como el tipo de iteraciones que intervienen en la construcción de las órbitas de billares poligonales y poliédricos, que son las iteraciones de una familia de isometrías, son un caso límite de las que sirven para la construcción de autosemejantes, en que se trata de iteraciones de una familia de semejanzas de razón menor que la unidad.

Cerramos el capítulo y el trabajo mostrando como todos los procesos descritos para la obtención de fractales se atienen a cierto esquema general, por lo que es de interés la pregunta de bajo que condiciones dicho esquema conduce a conjuntos de naturaleza fractal.

* * *

Quiero expresar aquí mi agradecimiento a todos quienes han hecho posible este trabajo. Ante todo a Miguel de Guzmán, cuyo encuentro en los cursos de doctorado ha sido un factor clave para decidirme a iniciar la actividad de investigación matemática, después de haber abandonado este campo durante algunos años. En la situación de desconexión en que me encontraba al reiniciar mis estudios, ha sido decisiva para mí la excelente acogida que Miguel de Guzmán me ha dispensado, poniéndome en contacto con las últimas tendencias de investigación matemática en el campo de teoría de iteraciones y teoría geométrica de la medida, temas con los que tan bien sintonizan mis propias aficiones. Por todo ello, la colaboración y el apoyo recibidos de Miguel de Guzmán, han significado para mí algo más importante que la posibilidad de realizar este trabajo, y ello ha sido la constatación de que es posible disfrutar haciendo matemáticas y que merece la pena hacerlo. Y precisamente uno de los requisitos indispensables para que esto sea posible, es poder realizar el trabajo de investigación en equipo. Nuevamente tengo aquí que agradecer a Miguel de Guzmán el marco que me ha ofrecido en el grupo de investigación sobre fractales, con participación de prestigiosos investigadores nacionales y extranjeros, y en el seminario de Teoría Geométrica de la Medida que dirige junto con Baldomero Rubio en la Universidad Complutense.

Quiero hacer constar mi especial agradecimiento a Baldomero Rubio por el aliento y colaboración que me ha brindado durante mis estudios de doctorado, y por las excelentes lecciones de análisis

que recibí de él durante mis estudios de licenciatura y sobre teoría geométrica de la medida en el curso de doctorado, y que fueron la mejor introducción a este campo que podía haber deseado.

Mi más entrañable gratitud a Miguel Angel Martín , Miguel Reyes, Ignacio Alvarez, Mara Izcue , María Teresa Carrillo, Domingo Casado, y a todos los compañeros y amigos que participan en el seminario de teoría geométrica de la medida, de los que he recibido mil sugerencias y con los que he contrastado mis ideas a lo largo de muchas horas de trabajo en equipo.

Mi agradecimiento a Martín Sánchez por su ayuda en los complicados diseños gráficos incluidos en el anexo al capítulo III, a Roberto Moriyón, a Francisco González Gascón, a Juan María López del Sa, al profesor Pertti Mattila, a Carles Simó y a todos cuantos me han ayudado con sus sugerencias , comentarios y orientaciones.

A handwritten signature in black ink, which appears to read 'Manuel Martín'. The signature is fluid and cursive, with a long horizontal stroke extending to the right.

Getafe, 27 de Mayo de 1988.

CAPITULO I

SOBRE LA CONJETURA DE LOS PENTAGONOS ENCAJADOS

DE SCHOENBERG

1.1. LA CONJETURA DE SCHOENBERG SOBRE LOS PENTAGONOS ENCAJADOS

1.1.1. La conjetura.

Sea K_1 un pentágono estrictamente convexo en el plano afín E_2 , donde por estrictamente convexo queremos decir que tres cualesquiera de sus vértices no son colineales. Sean A_i , $i \in Z_5$, los vértices de K_1 . Escribimos $A_i A_j$ para representar la recta que une esos dos puntos, y $(A_i A_j) \cdot (A_k A_l)$ representa el punto de corte de las dos rectas. Podemos definir los vértices de un nuevo pentágono estrictamente convexo (ver fig. 1.1) $K_2 \subset K_1$ por las siguientes fórmulas:

$$A_{2i} = (A_{i+1} A_{i+3}) \cdot (A_{i-1} A_{i-3}), \quad i \in Z_5$$

E inductivamente:

$$A_{j+1i} = (A_{ji+1} A_{ji+3}) \cdot (A_{ji-1} A_{ji-3}), \quad i \in Z_5$$

Obtenemos los vértices de una cadena de pentágonos estrictamente convexos:

$$K_{j+1} \subset \dots \subset K_3 \subset K_2 \subset K_1$$

Probaremos la siguiente conjetura de Schoenberg: $\bigcap_{i=1}^{\infty} K_i$ se reduce a un sólo punto. Esta conjetura fué formulada por I.J. Schoenberg en 1982 (ver ref [S1])

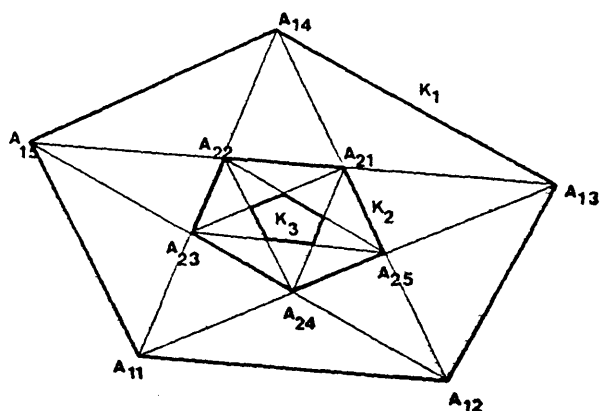


fig. 1.1.

1.1.2. Colineación asociada a K_1

Consideramos ahora el plano afín E_2 incluido en el plano proyectivo P_2 . Probaremos el siguiente lema:

Lema 1

Hay una colineación $f: P_2 \rightarrow P_2$ que aplica el vértice A_{ji} de K_j en el vértice A_{j+1i} de K_{j+1} , $j \in N$, $i \in Z_5$.

Para construir f , consideramos los dos conjuntos de cuatro puntos $\{A_{1i}\}$, $\{A_{2i}\}$, $1 \leq i \leq 4$. En estos conjuntos no hay subconjuntos de tres puntos colineales, de donde se deduce la existencia de una única colineación tal que:

$$f(A_{1i}) = A_{2i}, \quad 1 \leq i \leq 4$$

Probaremos en primer lugar que $f(A_{15})$ es precisamente A_{25} .

Para ello, sea $r_2 r_3 r_4 r_5$ el haz de rayos de vértice en A_{11} , donde $r_i = A_{11}A_{1i}$ (ver fig. 1.2). Consideramos también el haz de rayos $s_5 s_4 s_3 s_2$, con vértice en A_{21} , donde $s_i = A_{21}A_{2i}$. Estos dos haces cortan en los puntos $A_{12}, A_{24}, A_{23}, A_{15}$, que yacen en la línea $A_{12}A_{15}$, y por tanto son haces proyectivos, lo que indicaremos escribiendo:

$$s_5 s_4 s_3 s_2 \wedge r_2 r_3 r_4 r_5$$

Obviamente $r_2 r_3 r_4 r_5 \wedge r_5 r_4 r_3 r_2$, de donde:

$$s_5 s_4 s_3 s_2 \wedge r_5 r_4 r_3 r_2$$

Ahora buscamos la imagen del haz $r_5 r_4 r_3 r_2$ por f . Si $i = 2, 3, 4$, $f(r_i) = f(A_{11}A_{1i}) = A_{21}A_{2i} = s_i$. De esta manera, f aplica el haz

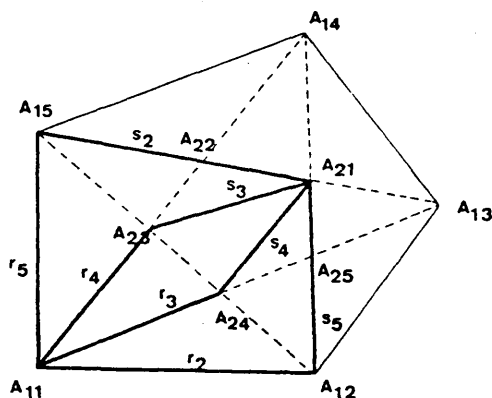


fig. 1.2

$r_5 r_4 r_3 r_2$ en el haz $f(r_5) s_4 s_3 s_2$, y, siendo f una colineación, podemos escribir $f(r_5) s_4 s_3 s_2 \wedge r_5 r_4 r_3 r_2$, y de aquí y de la anterior relación se obtiene que $f(r_5) s_4 s_3 s_2 \wedge s_5 s_4 s_3 s_2$. Esto prueba que $f(r_5) = s_5$. De manera similar, cambiando en nuestro razonamiento los haces $r_2 r_3 r_4 r_5$ y $s_5 s_4 s_3 s_2$ por los haces $m_3 m_2 m_1 m_5$, donde $m_i = A_{14} A_i$, y $t_5 t_1 t_2 t_3$, donde $t_i = A_{24} A_i$, se prueba que $f(m_5) = t_5$ y de aquí se obtiene finalmente:

$$f(A_{15}) = f(r_5 \cap m_5) = s_5 \cap t_5 = A_{25}$$

A continuación procedemos por inducción. Supongamos que el lema es cierto para $1 \leq j < n$. Tendremos:

$$\begin{aligned} f(A_{ni}) &= f((A_{n-1i+1} A_{n-1i+3}) \cdot (A_{n-1i-1} A_{n-1i-3})) = \\ &= (A_{ni+1} A_{ni+3}) \cdot (A_{ni-1} A_{ni-3}) = A_{n+1i}, \quad i \in \mathbb{Z}_5 \quad \blacksquare \end{aligned}$$

En lo que resta de este apartado llamaremos a la colineación obtenida así colineación asociada a K_1 , y la denotaremos con la letra f .

1.1.3. Propiedades de f

La colineación f aplica las rectas que contienen los lados de K_j en las que contienen los de K_{j+1} , $j \in \mathbb{N}$. Probaremos también la siguiente propiedad de f :

i) f aplica los lados de K_j en los de K_{j+1} y los puntos interiores de K_j en los de K_{j+1} , $j \in \mathbb{N}$.

Cuando hablamos de puntos interiores, nos referimos a los puntos interiores con respecto a la topología usual de $E_2 \subset P_2$. Identificamos E_2 con los puntos propios de P_2 , es decir, con aquellos que no pertenecen a la recta del infinito, recta que llamaremos E_1 . Sea $H = f^{-1}(E_1)$ y $F = P_2 - H$, y consideremos la restricción de f a F , $f|_F = g$. Esta aplicación g aplica F en E_2 , porque si $x \in F$, x no pertenece a H , y por tanto $f(x)$ no pertenece a E_1 . Seleccionando un sistema de referencia adecuado, probaremos que $g: F \rightarrow E_2$ es una aplicación continua e inyectiva. Elegimos un sistema de referencia en el que los puntos de P_2 tienen coordenadas homogéneas (x_0, x_1, x_2) , de tal forma que la recta impropia E_1 tenga ecuación $x_0 = 0$. Las ecuaciones de f serán.

$$\lambda x'_i = a_{i0} x_0 + a_{i1} x_1 + a_{i2} x_2, \quad i = 0, 1, 2$$

donde la matriz $A = (a_{ij})$, $0 \leq i \leq j$ es regular.

En este sistema de referencia todos los puntos de E_2 tienen coordenadas afines bien definidas:

$$\xi_i = x_i / x_0, \quad i = 1, 2$$

Las ecuaciones afines de g en F son:

$$\xi'_i = \frac{a_{i0} + a_{i1}\xi_1 + a_{i2}\xi_2}{a_{00} + a_{01}\xi_1 + a_{02}\xi_2}, \quad i = 1, 2$$

Estas ecuaciones prueban que g es inyectiva en F , y continua con respecto a la topología usual de F como subespacio de E_2 y P_2 . Así, si P y Q son puntos de E_2 y por $[PQ]$ representamos el segmento que une P con Q , siendo $[PQ] \subset F$, se tendrá $g([PQ]) = [g(P)g(Q)]$. Pero si suponemos que algún punto $X \in H$ es un punto interior de $[PQ]$, entonces $g([PQ]) = g(PQ) - [g(P)g(Q)]$, donde $g(PQ)$ es la recta imagen por g de la recta PQ , y recíprocamente el conjunto $PQ - [PQ]$ se aplica por g en el segmento $[g(P)g(Q)]$.

Sea $X \in H$, y supongamos que X pertenece a algún lado de K_1 , digamos a $[A_{11}A_{12}]$. No puede ser $X = A_{11}$ o $X = A_{12}$ porque si $X \in H$, $f(X) \in E_1$, y sabemos que los puntos $f(A_{11}) = A_{21}$ y $f(A_{12}) = A_{22}$ no están en E_1 . Luego X debe ser un punto interior a $[A_{11}A_{12}]$, y según se señalaba más arriba, $g([A_{11}A_{12}]) = A_{21}A_{22} - [A_{21}A_{22}]$. De esto concluimos que para algún punto $Y \in [A_{11}A_{12}]$ $g(Y) = A_{15}$ (ver fig. 1.3) y esto no es posible, porque el único punto que se aplica por g sobre A_{15} es $Z = (A_{11}A_{12}) \cdot (A_{13}A_{12})$, y Z no pertenece a $[A_{11}A_{12}]$. Así hemos probado que no es posible que X pertenezca a la frontera de K_1 . Además $K_1 \subset F$, porque si $K_1 \cap H \neq \emptyset$, siendo K_1 convexo habría puntos de H en la frontera de K_1 .

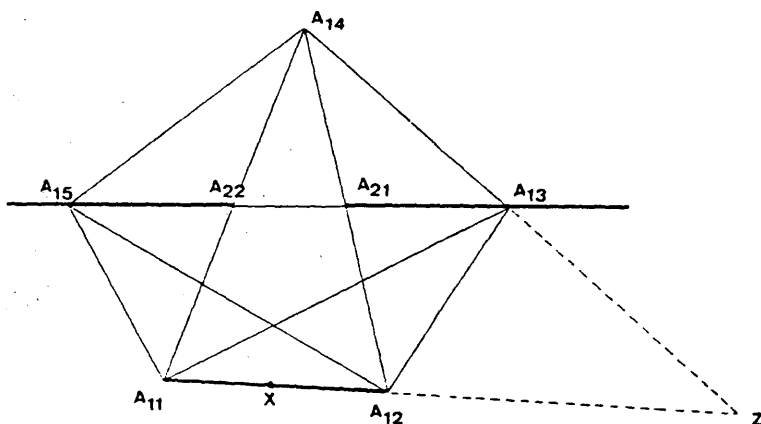


fig. 1.3

Ahora está claro que si $A, B \in K_n$, $n \in \mathbb{N}$, $g([AB])$ es un segmento con extremos $g(A)$ y $g(B)$. Por consiguiente cada lado de K_n se aplica en el lado respectivo de K_{n+1} y $g(\text{Fr}(K_n)) = \text{Fr}(K_{n+1})$. El conjunto $g(K_n)$ es un conjunto convexo con la misma frontera que K_{n+1} y por tanto coincide con K_{n+1} . ■

De estos resultados podemos obtener fácilmente el siguiente lema:

Lema 2

La colineación f asociada a K_1 tiene las siguientes propiedades:

- a) Tiene un punto fijo $P \in \bigcap K_n$
- b) P es el único punto fijo de f en K_1
- c) P está en todas las rectas invariantes por f que cortan a K_1

La existencia de P es un resultado del teorema de Brouwer aplicado a la restricción g de f a K_1 , que es, como hemos probado, una aplicación continua en K_1 tal que $g(K_1) = f(K_1) = K_2 \subset K_1$. Así g tiene un punto fijo $P \in K_1$ y $f(P) = g(P) = P$. Si suponemos que $P \in K_n$, $f(P) = P \in f(K_n) = K_{n+1}$, y por inducción, concluimos a).

Sea ahora Q un punto en la frontera de K_1 , como en la figura 1.4, y $[PQ]$ el segmento que une el punto fijo P con Q . Si $Q \in [A_{11}, A_{12}]$, $Q' = f(Q) \in [A_{21}, A_{22}]$ y $g([PQ]) = [PQ']$, y como $[PQ] \cap [PQ'] = P$, concluimos que el único punto fijo en $[PQ]$ es P . Si Q varía a lo largo de la frontera de K_1 , los segmentos $[PQ]$ cubren K_1 , y esto prueba que P es el único punto fijo de f en K_1 .

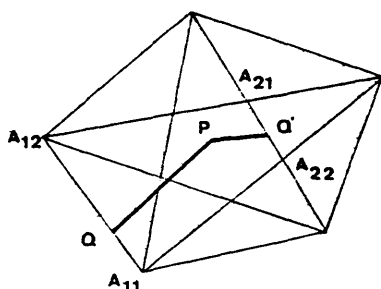


fig. 1.4

Finalmente para probar c), supongamos que r es una recta tal que $f(r) = r$, y $r \cap K_1 \neq \emptyset$. Entonces $f(r \cap K_1) = r \cap f(K_1) = r \cap K_2 \subset r \cap K_1$. Aplicando el teorema de Brouwer a la restricción de f a $r \cap K_1$, obtenemos la existencia en dicho segmento de un punto fijo de f . Este punto debe ser P , por la propiedad b), y c) queda demostrada. ■

En adelante, siempre denotaremos por P el único punto fijo de la colineación f en K_1 .

1.1.4. Clasificación de f

En primer lugar excluirémos una posibilidad que haría trivial el resto de la demostración: El ejemplo mostrado en la figura 5 prueba que, en general f no es una aplicación contractiva. En efecto, $d(A_{11}, A_{12}) < d(f(A_{11}), f(A_{12})) = d(A_{21}, A_{22})$. Según ha

señalado C. Simó, tampoco puede encontrarse en general una constante $C < 1$ tal que, si $|K_n|$ representa el diámetro de K_n , sea $|K_{n+1}| < C \cdot |K_n|$, lo que permitiría también una fácil prueba de la conjetura. En efecto, aplastando convenientemente el pentágono K_1 de la figura 1.5, y achatándolo por el vértice A_{12} , puede conseguirse que el diámetro de K_2 sea tan parecido como se desee al de K_1 .

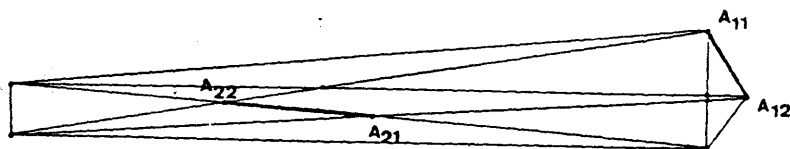


fig. 1.5

De esta forma, para conseguir la demostración de la conjetura, debemos pasar a estudiar las iteraciones de la colineación f , para lo cual es necesario clasificar los distintos casos que se pueden presentar. Haremos partir esta clasificación de una propiedad geométrica de f : La existencia de alguna recta invariante por f que no pase por el punto fijo P .

Si tal recta existe, eligiendo un sistema de referencia en el que la recta tenga ecuaciones afines $x_0 = 0$, las ecuaciones de f en dicho sistema serán las de una afinidad, y puede demostrarse la conjetura basándose en sus propiedades.

Si todas las rectas invariantes por f pasan por P , la raíz de la ecuación característica asociada a P debe ser doble o triple, ya que (ver ref. [S3]) un punto fijo asociado a una raíz simple no puede pertenecer a la recta invariante asociada a la misma raíz. Esto limita las posibles formas canónicas de f a dos, y estudiando las iteraciones de las mismas, se obtiene el teorema.

Debemos pues analizar los tres siguientes casos:

1.1.5. Caso 1

Suponemos aquí que existe r tal que $f(r) = r$ y P no pertenece a r . Expresamos las ecuaciones de f en un sistema de referencia en que r tenga ecuación $x_0 = 0$. En tal sistema de referencia las ecuaciones de f son las de una afinidad (invarianza de la recta impropia), y las coordenadas afines de los puntos de K_1 son finitas en dicho sistema de referencia, porque sabemos por c) del lema 2 que $r \cap K_1 = \emptyset$. Si S_n es la superficie de K_n en este sistema de referencia, S_n es un número real finito y positivo, y existe un módulo de afinidad $t \in \mathbb{R}$, tal que $S_{n+1} = t \cdot S_n$. Además, como las inclusiones de los pentágonos son estrictas, t es menor que 1, y por tanto $S_n \rightarrow 0$, de donde $\cap K_n$ es un cuerpo convexo de superficie nula, es decir, un segmento o un punto. Probaremos la imposibilidad de la primera

hipótesis, y de ello se deducirá, en este caso, la veracidad de la conjetura.

Supongamos $\cap K_n = [BC]$. Sería $f([BC]) = f(\cap K_n) = \cap K_n = [BC]$. La frontera de $[BC]$ se aplica en ella misma, por la continuidad e inyectividad de $g = f/K_1$, de manera que bien $f(B) = B$ y $f(C) = C$, en contradicción con b) del lema 2, o bien $f(B) = C$ y $f(C) = B$, en cuyo caso la restricción de f a la recta AB , que es una proyectividad, sería una involución. Pero si $Q \in Fr(K_1) \cap AB$, $f^2(Q) \in Fr(K_3)$, luego $f^2(Q) \neq Q$, y esto prueba que f/AB no puede ser una involución. ■

1.1.6. Caso 2.

Suponemos ahora que P pertenece a cada recta r tal que $f(r) = r$, y que P está asociado a una raíz triple de la ecuación característica de f . Las forma canónica de las matriz de f y su potencia n -ésima, serán respectivamente:

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & n & \begin{bmatrix} n \\ 2 \end{bmatrix} \\ 0 & 1 & n \\ 0 & 0 & 1 \end{bmatrix}$$

Hemos supuesto que la raíz triple asociada a P es 1 ya que siendo f una colineación, podríamos multiplicarla por una constante real no nula si fuera necesario.

El caso que estamos analizando es el único caso en el que f tiene raíz triple, porque las otras dos posibles formas canónicas en que f tiene una raíz triple, corresponden a la identidad y a una elación con toda una recta de puntos fijos pasando por P , tienen infinitos puntos fijos sobre K_1 , en contradicción con b) del lema 2.

Sea pues f la colineación con la forma canónica escrita más arriba. $P(1,0,0)$ es su único punto fijo, y es por tanto el mismo P obtenido por el lema 2. Sea $X(x_0, x_1, x_2)$ un punto general de P_2 , y $f^n(X) = X(n)(x_0(n), x_1(n), x_2(n))$. Tenemos:

$$\begin{aligned} \lambda x_0(n) &= x_0 + x_1 n + x_2 \begin{bmatrix} n \\ 2 \end{bmatrix} \\ \lambda x_1(n) &= x_1 + x_2 n \\ \lambda x_2(n) &= x_2 \end{aligned}$$

Como $x_0 = x_1 = x_2 = 0$ no se verifica para ningún punto en P_2 , $x_0(n)$ no es el polinomio nulo, de forma que para algún n_0 , $n > n_0 \Rightarrow x_0(n) \neq 0$, y las coordenadas afines de $X(n)$ están bien definidas: Estas coordenadas son $\xi_i(n) = x_i(n)/x_0(n)$, $i = 1, 2$. En estas funciones racionales, fijado el punto X , el polinomio del denomina-

donde tiene el grado mayor que el del numerador, con el único caso excepcional en que $x_1 = x_2 = 0$, y en este caso X es precisamente $X(1,0,0) = P$, de donde $X(n) = P$ y por tanto $X(n) \rightarrow P$. Pero esta relación se verifica también en el caso general, ya que según la observación hecha sobre los grados de los polinomios en el caso general, $X(n) \rightarrow (0,0)$ en coordenadas afines, y $(0,0)$ es precisamente la expresión de $P(1,0,0)$ en coordenadas afines. Se ha probado que para cualquier $X \in P_2$, $X(n) \rightarrow 0$. De ello se deducirá fácilmente en 1.1.8 la verificación de la conjetura en este caso.

En este caso, las iteraciones f convergen hacia el punto fijo atractivo P para cualquier punto de P_2 . No obstante, es interesante señalar que la aplicación f no es contractiva.

1.1.7. Caso 3.

Suponemos ahora que P es el punto fijo asociado a una raíz doble de la ecuación característica de f , última posibilidad que nos queda por explorar.

Las formas canónicas de f y f^n son ahora, respectivamente:

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & p \end{bmatrix} \quad \begin{bmatrix} 1 & n & 0 \\ 0 & 1 & 0 \\ 0 & 0 & p^n \end{bmatrix}$$

Donde $p \neq 1$, ya que el caso $p = 1$ corresponde a una raíz triple, concretamente al caso de la elación, que ya hemos descartado en el punto anterior.

Los puntos fijos para f son $P(1,0,0) \in K_1$ — ya que este es el punto fijo asociado a la raíz doble —, y $Q(0,0,1)$ que no pertenece a K_1 , según b) del lema 2. Si con la misma notación del caso anterior es $X \in P_2$ y $X(n) = f^n(X)$, tendremos:

$$\lambda x_0(n) = x_0 + x_1 n$$

$$\lambda x_1(n) = x_1$$

$$\lambda x_2(n) = x_2 p^n$$

Si excluimos el caso $X = Q$, el polinomio $x_0(n)$ no es el polinomio nulo, y no vale cero para valores de n superiores a cierto n_0 . Para esos valores de n , las coordenadas afines de $X(n)$ están bien definidas.

Nos proponemos probar en primer lugar que $|p| < 1$. La recta r de ecuación $x_1 = 0$ que contiene los puntos P y Q es invariante por f . Su ecuación afín es $\xi_1 = 0$, y la restricción de f^n a r tiene ecuación $\xi_2(n) = p^n \xi_2$. Sea $x \in r \cap K_1$ y $|p| > 1$. La sucesión

$\xi_2(n)$ no estaría acotada. Pero esto no es posible, porque el punto $X(n) \in r \cap K_1$ y tiene coordenadas afines $(0, \xi_2(n))$, y podemos encontrar una cota para las coordenadas afines de los puntos de $r \cap K_1$, puesto que el único punto impropio de la recta r es Q , y Q no es adherente a $r \cap K_1$. Esta contradicción prueba que $|p| \leq 1$. La posibilidad $|p| = 1$, se reduce a $p = -1$, ya que sabemos que $p \neq 1$. Y si $p = -1$, la restricción de f a r sería una involución, lo que sabemos que no puede ser, según la argumentación al respecto presentada en el caso 1. Queda como posibilidad única $|p| < 1$.

Sea ahora $|p| < 1$ y $X \in P_2$, $X \neq Q$. Demostraremos bajo estas hipótesis que $X(n) \rightarrow P$. Basta observar las ecuaciones afines de f . Se tiene $\xi_1(n) \rightarrow 0$, $\xi_2(n) \rightarrow 0$, de donde $X(n) \rightarrow (0, 0)$, siendo este último punto la expresión de P en coordenadas afines.

Vemos que las iteraciones de f convergen hacia el punto fijo atractivo P , excepto en el caso de Q , que es un punto fijo repulsivo, pero que no pertenece a K_1 .

1.1.8. Conclusión

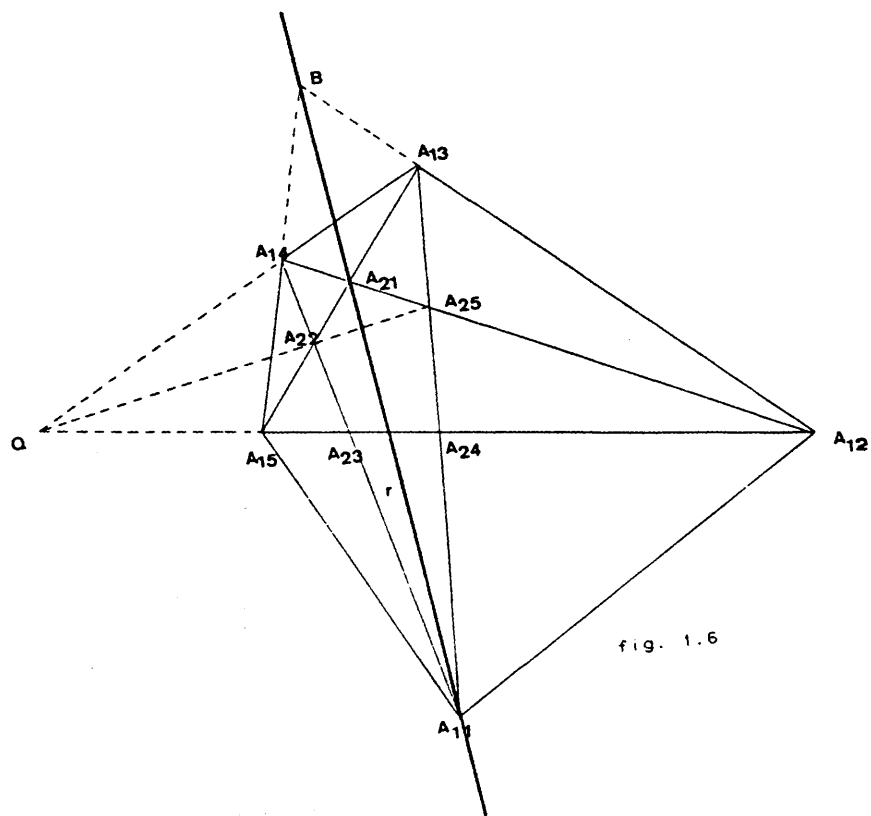
Concluimos la prueba de la conjetura en los casos 2 y 3, observando que en ambos casos, si $X \in K_1$, $X(n) \rightarrow P$. Fijemos una bola $B_\epsilon(P)$ centrada en P y de radio arbitrario ϵ . Para un n suficientemente grande, $A_{ij}(n) \in B_\epsilon(P)$, $i \in \mathbb{Z}_5$. De aquí se deduce que $K_n \subset B_\epsilon(P)$, y por tanto $\bigcap_j K_j \subset B_\epsilon(P)$, y de ello concluimos que $\bigcap_j K_j = P$. ■

1.1.9. Construcción gráfica de P

Si conocemos una recta invariante por f , o un punto fijo de f , se puede construir P con regla y compás. Para ello, obtenemos todos los puntos fijos de f , y seleccionamos el único que pertenezca a K_1 . La obtención de los puntos fijos de f es, en general, un problema de tercer grado. Pero si se conoce un punto fijo o una recta invariante por f , el problema resulta de segundo grado, y puede ser resuelto con regla y compás.

Presentamos en la figura 1.6 un ejemplo de este tipo, donde los puntos $B = (A_{12}A_{13}) \cdot (A_{14}A_{15})$, A_{11} y A_{21} son colineales. La línea r que los une debe ser invariante, porque $f(r) = f(BA_{11}) = f(B)A_{21} = A_{11}A_{21} = r$. Podemos obtener los puntos fijos de f/r con regla y compás mediante una construcción bien conocida (ver ref. [S1]). Es interesante señalar que aplicando el teorema de Desargues a los triángulos $A_{13}A_{12}A_{25}$ y $A_{14}A_{15}A_{22}$, se deduce la concurrencia de las rectas $A_{13}A_{14}$, $A_{12}A_{15}$ y $A_{25}A_{22}$ en el punto Q , que es un punto fijo, porque $f(Q) = f((A_{13}A_{14}) \cdot (A_{12}A_{15})) = (A_{23}A_{24}) \cdot (A_{22}A_{25}) = Q$. Recíprocamente, de la concurrencia de las tres líneas en Q , hubiéramos podido deducir que Q es fijo y que los puntos B, A_{11} y A_{21} están alineados en una recta r invariante.

Otro ejemplo característico, es aquel en que K_1 es un pentágono regular. En este caso la línea del infinito tiene todos sus puntos fijos, f es una aplicación afín, y P es el centro del pentágono regular.



CAPITULO II

ISOMETRIAS LOCALES BIYECTIVAS EN \mathbb{R}

2.0. INTRODUCCION.

El presente capítulo está dedicado íntegramente al estudio de las isometrías locales en los números reales, centrándose en particular en las propiedades de minimalidad de las mismas.

El plan del capítulo es el siguiente. En el apartado 2.1, se da la definición de intercambio de intervalos, y una definición de i.l.b., que parte de propiedades locales, pasándose en el apartado 2.2 a obtener las primeras propiedades globales de las i.l.b. En el apartado 2.3 se consigue clasificar el dominio de las i.l.b. en una parte periódica y una parte acíclica, de forma que bajo cierta hipótesis complementaria a la definición, que luego desempeña también un importante papel en las propiedades de las órbitas, la restricción de la i.l.b. a la parte acíclica sigue siendo i.l.b. En el apartado 2.4 se pasa a estudiar las i.l.b. acíclicas, encontrándose que su estudio puede ser reducido al caso en que el dominio conste de un número finito de componentes conexas de diámetro positivo, llamándose r.i. a las i.l.b. de este tipo. En el apartado 2.5 se obtienen algunas propiedades de las r.i., y en el apartado 6, que es el núcleo central del capítulo, se estudian las propiedades de las órbitas de las r.i. acíclicas, obteniéndose bajo hipótesis complementarias los teoremas fundamentales, que son el de recurrencia, el de estructura, y el de minimalidad. Finalmente en el apartado 2.7 se consigue una flexibilización de la definición de r.i., útil en las aplicaciones, y se extienden los teoremas fundamentales a r.i. que no satisfacen las hipótesis complementarias requeridas en el apartado 2.6.

2.1. DEFINICIONES.

2.1.1. Definición de intercambio de intervalos (M. Keane 1975)

Sea $X = [0,1[$ y $n > 1$ un entero. Para cada vector de probabilidad :

$$\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$$

con $\alpha_i > 0$, para cada $1 \leq i \leq n$ escribimos :

$$\beta_0 = 0; \beta_i = \sum_{j=1}^i \alpha_j;$$

$$X_i = [\beta_{i-1}, \beta_i[$$

Sea τ una permutación de los símbolos $1, 2, 3, \dots, n$. Entonces

$$\alpha^\tau = (\alpha_{\tau^{-1}(1)}, \dots, \alpha_{\tau^{-1}(n)})$$

Es un vector de probabilidad con componentes positivas, y podemos

formar los correspondientes β_i^T y X_i^T , $1 \leq i \leq n$.

Definimos ahora $T : X \rightarrow X$ poniendo :

$$Tx = x - \beta_{i-1} + \beta_i^T - 1$$

para cada $x \in X_i$ y cada $1 \leq i \leq n$. T aplica isométricamente cada X_i en el correspondiente intervalo X_i^T .

Llamamos a T el intercambio de intervalos (α, τ) .

Las transformaciones así definidas no son otras que las isometrías a trozos del intervalo que preservan el orden. Los puntos de discontinuidad son los extremos de los intervalos X_i , es decir los β_i , pero se conserva la continuidad a la derecha en dichos puntos.

2.1.1 Condición de minimalidad.

Si T el intercambio de intervalos (α, τ) , escribimos, para cada $x \in X$:

$$O^+(x) = \{ T^t x : t \geq 1 \}$$

$$O^-(x) = \{ T^t x : t \leq 0 \}$$

$$O(x) = O^+(x) \cup O^-(x)$$

Decimos que T satisface la condición de minimalidad si :

M 1) T es aperiódica, es decir cada $O(x)$ es infinito

M 2) Si F es una unión finita de intervalos semiabiertos cuyos puntos extremos pertenecen al conjunto numerable :

$$D^F = \bigcup_{i=0}^{n-1} O(\beta_i) \cup 1$$

Entonces $TF = F$ implica $F = X$ o $F = \emptyset$

El interés de la condición de minimalidad reside en el siguiente resultado :

Teorema (Keane 1975) : T satisface la condición de minimalidad si y sólo si $O(x)$ es denso en X para cada $x \in X$.

2.1.2. Definición de isometría local

Sea $S \subset \mathbb{R}$ un conjunto localmente conexo y acotado. $f: S \rightarrow S$ es una isometría local biyectiva si es una biyección con un conjunto finito $X(f)$ de puntos de discontinuidad - que llamaremos puntos de cambio - y tal que si $x \in S$ es un punto no aislado, existe un intervalo $I(x)$ de diámetro positivo tal que:

- $x \in I(x) \subset S$
- $f/I(x)$ es una isometría

Las isometrías locales son biyecciones que transportan de forma isométrica cada punto no aislado junto con todo un intervalo que lo contiene.

Para indicar que f es una isometría local biyectiva escribiremos de aquí en adelante : f es una i.l.b.

Si para cada intervalo $I(x)$ tal que $f/I(x)$ es isometría, se sabe además que la isometría es creciente, entonces diremos que f es orientable.

En la siguiente figura se representa un ejemplo de isometría local.

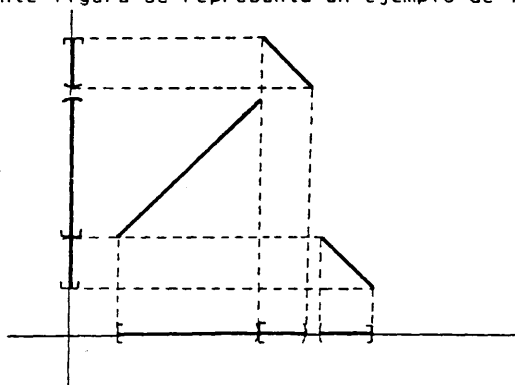


fig. 2.1

2.2. PROPIEDADES GLOBALES DE LAS ISOMETRIAS LOCALES BIYECTIVAS

En este apartado se estudian las consecuencias en el comportamiento global de las i.l.b. de su definición, que se basa en propiedades de carácter esencialmente local. Antes de abordar su comportamiento respecto a las iteraciones, que es el objetivo del capítulo, en este y los siguientes apartados pondremos las bases que nos permitirán una clasificación previa de las i.l.b. necesaria para el estudio posterior de sus iteraciones.

2.2.1 Conservación de la medida

Sea $f: S \rightarrow S$ una i.l.b. Si $A \subset S$ es medible Lebesgue, entonces se verifica que la medida de A , $L(A)$, coincide con $L(f^{-1}(A))$.

El subconjunto K de los puntos aislados de S es numerable, y basta por tanto probar que si un conjunto medible A está incluido en $S-K$, se verifica la condición expuesta. Pero los puntos del conjunto $X(f)$ dividen las componentes conexas de $S-K$, que son intervalos de diámetro positivo, en intervalos en cuya restricción f es continua. Es fácil ver que la restricción de f a cada uno de esos intervalos, ha de ser una isometría de la recta, ya que si dos puntos de uno de tales intervalos se vieran transportados cada uno junto con todo un subintervalo, de forma isométrica, por dos isometrías diferentes, debería haber entre ambos algún punto de discontinuidad. De este hecho se deduce inmediatamente el teorema. ■

2.2.2. Componentes isométricas

El razonamiento anterior, prueba además que f es una isometría a trozos. La siguiente definición, es de uso constante.

Llamaremos intervalo isométrico a todo intervalo I tal que $f|I$ es isometría. Las componentes isométricas se definen como intervalos isométricos maximales.

2.2.3 Propiedades de las componentes isométricas.

Cada punto está contenido en una componente isométrica, ya que la definición de i.l.b. garantiza que no es vacía la familia de los intervalos isométricos que contienen un punto y la componente isométrica de un punto x , que denotaremos por $J(x)$, se obtiene como unión de todos los intervalos de dicha familia. Si el punto x es aislado, $J(x)$ se reduce al punto. En caso contrario $J(x)$ es un intervalo de diámetro positivo.

De esta forma las componentes isométricas forman una partición de S en intervalos isométricos. Dicha partición es más fina que la partición en componentes conexas.

Los extremos de las componentes isométricas no pueden ser puntos del conjunto $\text{Int}(S)-X(f)$ porque si lo fuesen, según lo expuesto más arriba respecto de los puntos de dicho conjunto, los intervalos maximales podrían ampliarse. De ello se deduce que dichos extremos son puntos frontera de S o puntos de $X(f)$. De hecho, los extremos de las componentes isométricas que están contenidos en S coinciden con la unión de los puntos de cambio y los puntos frontera de S incluidos en S . Por otra parte es claro que los puntos de cambio deben ser interiores a S . Todo lo dicho se expresa en las fórmulas:

$$i) X(f) \subset \text{Int}(S)$$

$$ii) \bigcup_{x \in S} \text{Fr}(J(x)) \subset \text{Fr}(S) \cup X(f)$$

$$iii) \bigcup_{x \in S} (\text{Fr}(J(x)) \cap S) = (\text{Fr}(S) \cap S) \cup X(f)$$

Se ha observado más arriba que la partición del dominio S de una i.l.b. en componentes isométricas es más fina que la partición en componentes conexas. Cada componente conexa está dividida en componentes isométricas. De hecho son los puntos de cambio los que separan entre sí las diferentes componentes isométricas de una componente conexa, de manera que si una componente conexa no contiene puntos de cambio, entonces coincide con una componente isométrica. Si designamos por N_C el número de componentes conexas, por N_I el de componentes isométricas y por N_X el de puntos de cambio, se prueba fácilmente que:

$$IV) N_I = N_C + N_X$$

Manteniendo validez esta relación incluso si N_C o N_I son infinitos.

2.3. PARTE PERIÓDICA DE UNA I.L.B.

En este apartado nos proponemos clasificar los puntos del dominio de una i.l.b. en dos tipos: aquellos que son periódicos, y los que no lo son. Se probará que bajo ciertas condiciones, la restricción de la i.l.b. a la parte no periódica es una i.l.b. Las condiciones necesarias para ello motivan la introducción de los conceptos de i.l.b. regular y semirregular. Se analiza completamente la estructura de la parte periódica, agrupando sus puntos en ciclos y haces de ciclos.

2.3.1. Definiciones

Sea $f: S \rightarrow S$ una aplicación cualquiera. Recordamos los siguientes conceptos de uso habitual en teoría de iteraciones:

i) Ciclo de orden k o k -ciclo, es una k -tupla de elementos de S diferentes dos a dos, $\{x_i\}$, $1 \leq i \leq k$ tales que $f(x_i) = x_{i+1}$, $1 \leq i < k$ y $f(x_k) = x_1$.

ii) f es periódica en $x \in S$ si x pertenece a algún ciclo, es decir, si para cierto $k \in \mathbb{N}$ es $f^k(x) = x$. Los elementos de un k -ciclo se llaman cíclicos o periódicos de orden k . Se dice que f es periódica por puntos si es periódica en cada $x \in S$. Se dice que f es periódica en S si para algún k es $f^k(x) = x$ para cada $x \in S$.

2.3.2.

Es fácil obtener el siguiente resultado: Si $f: S \rightarrow S$ es una i.l.b. y C es un conjunto S -cerrado tal que $f(C) = C$, la restricción de f a $S-C$ es también una i.l.b.

En particular, como los ciclos verifican las citadas condiciones, podemos quitar un ciclo del dominio, y la restricción de f al

conjunto restante será i.l.b. Desgraciadamente no se puede proceder a quitar todos los ciclos de uno en uno hasta conseguir una i.l.b. sin ciclos, ya que los ciclos se presentan por lo general en cantidad no finita e incluso no numerable. Utilizando el resultado de más arriba, podremos no obstante proceder a retirar los ciclos por grupos que llamaremos haces de ciclos.

Sin embargo lo que sí podremos suponer de ahora en adelante es que no hay puntos cíclicos en $X(f)$. En efecto, si los hubiera, como $X(f)$ es un conjunto finito, los podríamos retirar uno a uno junto con todo su ciclo aplicando el resultado mencionado en este punto un número finito de veces.

2.3.3. Intervalos k -isométricos y ω -isométricos.

Sea f una i.l.b. en S . Un intervalo J incluido en S es k -isométrico si verifica:

$$f^i(J) \cap X(f) = \emptyset \quad 0 \leq i < k$$

Se considera $f^0 = \text{id}$.

Un intervalo 1-isométrico es simplemente un intervalo isométrico.

La propiedad que se deriva inmediatamente de esta definición es la siguiente:

Si J es k -isométrico, f^i/J es isometría si $1 \leq i \leq k$.

Finalmente llamamos a un intervalo J ω -isométrico si es k -isométrico para cualquier $k \in \mathbb{N}$. Si J verifica tal condición, se cumple que f^k/J es isometría para cualquier $k \in \mathbb{N}$.

2.3.4. Componentes cíclicas y haces de ciclos

Seguimos suponiendo - ver 2.3.2. - que en $X(f)$ no hay puntos cíclicos. Entonces, se puede probar que si $x \in S$ es un punto cíclico de orden k , hay un intervalo $I(x)$ compuesto de puntos cíclicos de orden $2k$ que verifica $x \in I(x) \subset S$, e $I(x)$ es S -abierto. Tal intervalo se obtiene tomando una S -bola centrada en x y de radio lo suficientemente pequeño para que no haya en ella puntos del conjunto:

$$D = \bigcup_{i=0}^{k-1} f^{-i}(X(f))$$

Esto es posible debido a que D es finito y a que x no puede estar en D si hemos supuesto que en $X(f)$ no hay puntos cíclicos. En estas condiciones $I(x)$ es k -isométrico. Descartando el caso trivial de que x sea un punto aislado, en cuyo caso $I(x) = x$ verifica lo requerido, x puede ser o bien extremo de una componente conexa de S , o bien interior a S . En el primer caso $f^k/I(x)$ es una isometría que aplica x en x y no puede ser decreciente o x no sería un punto frontera de S . De ello se deduce que dicha aplicación es una

isometría creciente en $I(x)$, es decir, la identidad. Si x es interior a S , podemos escoger $I(x)$ lo suficientemente pequeño para que esté incluido en su componente conexa, es decir para que sea una bola centrada en x (en el caso general era una S -bola). Entonces $f^k/I(x)$ aplica isométricamente $I(x)$ sobre $I(x)$. En el caso de que fuera una isometría decreciente, $f^{2k}/I(x)$ sería la identidad. ■

Definimos la *componente cíclica* de un punto k -cíclico x como el intervalo $P(x)$ maximal que contiene a x y en el que $f^k/P(x)$ es isometría. Tal intervalo es no vacío, según lo razonado más arriba. El mismo razonamiento, junto al carácter maximal de $P(x)$ prueba que todos los puntos de dicho intervalo son S -interiores a él y por tanto se trata de un intervalo S -abierto.

Además sabemos que está todo él formado de puntos k -cíclicos o $2k$ -cíclicos.

Definimos *haz cíclico* de un punto k -cíclico x como el conjunto $H(x)$ unión de todos los ciclos de los puntos que están en la componente cíclica de x . Es claro que $H(x)$ se obtiene también:

$$H(x) = \bigcup_{i=0}^{k-1} P(f^i(x))$$

Finalmente observaremos que la componente cíclica de un elemento cíclico es ω -isométrica. Más adelante se probará que la única clase posible de intervalos ω -isométricos de diámetro positivo que puede haber son los subintervalos de una componente cíclica.

2.3.5. Haces de Ciclos S-cerrados.

En este punto vamos a caracterizar los haces de ciclos S -cerrados. Probaremos que el conjunto C de todos ellos es S -cerrado y por tanto a él se puede aplicar lo dicho en 3.3.2. y podemos retirar así todos los haces de ciclos S -cerrados de S , siendo la restricción de f al conjunto restante una i.l.b. Supondremos, como en los anteriores puntos, que en $X(f)$ no hay puntos cíclicos.

Supongamos que el haz de ciclos de cierto punto x k -cíclico, $H(x)$, es tal que en su adherencia relativa a S no hay puntos de cambio. De esta forma los extremos de las componentes cíclicas que forman $H(x)$ no pueden ser puntos de cambio ni tampoco pueden ser puntos cuya imagen por una iteración de f sea un punto de cambio ya que en tal caso la imagen de la componente cíclica por la correspondiente iteración tendría, por continuidad, un extremo en $X(f)$. Esto prueba que las iteraciones de f son siempre continuas en los extremos de las componentes cíclicas de $H(x)$ si dichos extremos pertenecen a S . Dichos extremos deben ser entonces cíclicos y por la maximalidad de las componentes cíclicas, deben pertenecer a ellas. Esto prueba que las componentes cíclicas de $H(x)$ son S -cerradas. De aquí se deduce que $H(x)$ es unión finita de S -cerrados, y por tanto S -cerrado.

Por otra parte, si $H(x)$ es S -cerrado, en su adherencia relativa a

S , que coincide con $H(x)$, no puede haber puntos de cambio, ya que estos serían cíclicos. Hemos probado así:

$$H(x) \text{ es } S\text{-cerrado} \Leftrightarrow H(X) \cap S \cap X(f) = \emptyset$$

Finalmente probaremos que *la unión C de todos los haces de ciclos S -cerrados es S -cerrada*. Como cada haz de ciclos S -cerrado se obtiene como unión de componentes conexas de S , la unión de todos ellos también se pone como unión de componentes conexas de S . Naturalmente también $S-C$ es unión de componentes conexas. Como S es localmente conexo, cada componente conexa suya es S -abierto, por lo que $S-C$ es S -abierto y C es S -cerrado. ■

2.3.6. Isometrías locales regulares y semirregulares

Debemos ahora proceder a retirar los haces de ciclos que no son S -cerrados. Para que ello sea posible es preciso introducir una hipótesis complementaria: la de regularidad, que mas tarde juega un importante papel en el estudio de las iteraciones de f . Veremos que también puede utilizarse otra hipótesis mas débil, la de semirregularidad.

Sea $f: S \rightarrow S$ una i.l.b. con puntos de cambio en $X(f)$ tal que $f^n(X(f)) \cap X(f) = \emptyset$ para todo $n \in \mathbb{N}$. Entonces decimos que f es regular. La condición de regularidad se puede también expresar así: $f^{-n}(X(f)) \cap X(f) = \emptyset$ para todo $n \in \mathbb{N}$.

La propiedad fundamental de las i.l.b. regulares es la siguiente:

Si f es una i.l.b. regular entonces todas las iteraciones de f también son i.l.b. y por añadidura regulares.

Sea en efecto f una i.l.b. regular en S , y consideremos una iterada cualquiera suya, $g = f^K$, $K \in \mathbb{N}$. Consideremos también el conjunto:

$$D = \bigcup_{i=0}^{K-1} f^{-i}(X(f))$$

Si x no está en D , tomando una S -bola centrada en x y de radio suficientemente pequeño para no incluir puntos de D , obtenemos un intervalo K -isométrico, que es de diámetro positivo si x no es aislado. Dicho intervalo es el $I(x)$ requerido en la definición de i.l.b. para que g lo sea. Esta propiedad la tienen los puntos x que no están en D incluso si f no es regular.

Si x está en D , la regularidad hace que la distancia de x a $D - \{x\}$ sea positiva, y nuevamente podemos encontrar el intervalo requerido. Como por otra parte los puntos de cambio de g , $X(g)$, están incluidos en el conjunto finito D , resulta que g es i.l.b. en S . ■

La condición de regularidad, suficiente para que las iteradas de f sean i.l.b. en S , no es sin embargo necesaria. Se puede substituir por otra más débil, la exigida en la siguiente definición:

Una i.l.b. en S es semirregular si para todo punto no aislado $x \in S$ y para todo $k \in \mathbb{N}$, podemos encontrar un intervalo k -isométrico $I(x)$ de diámetro positivo tal que $x \in I(x)$.

Obsérvese que esto es mas fuerte que encontrar un intervalo isométrico para la iterada k -ésima. Exige que también sea isométrico para todas las anteriores. Por tanto si f es semirregular, todas las f^k son también isometrías locales.

2.3.7. Resultado final.

Los haces de ciclos restantes, una vez retirados de S los haces S -cerrados, son aquellos cuya S -adherencia corta a $X(f)$. No pueden ser sino un número finito, porque cada punto de $X(f)$, que suponemos como siempre no periódico, como no puede pertenecer al interior de las componentes cíclicas de un haz de ciclos, puede ser extremo de a lo sumo dos componentes cíclicas y por tanto de dos haces cíclicos, el número de estos puede duplicar, como máximo, al de puntos de $X(f)$. De aquí se desprende que si f es regular o semirregular, los podremos retirar uno a uno, aplicando la siguiente proposición:

i) Si f es regular en S , $H(x)$ es el haz de ciclos de un elemento k -cíclico x y $S' = S - H(x)$, $g = f/S'$ es una i.l.b. en S' .

En primer lugar se obtiene fácilmente la conexión local de S' . También es claro que $X(g)$ es finito. Si z es un punto del S' -interior de S' , tomando un intervalo $I(z)$ en el que f actúe isométricamente y con diámetro lo suficientemente pequeño para que esté incluido en S' , se tiene el intervalo requerido en la definición de i.l.b. para que g lo sea. Si $z \in S'$ no es un punto S' -interior de S' , es un punto frontera de alguna componente cíclica P de $H(x)$. Entonces, para algún $i \in \mathbb{N}$ $f^i(z) \in X(f)$ o de lo contrario por continuidad y por la maximalidad de P , $z \in P$, y sabemos que esto no es posible por ser las componentes cíclicas S -abiertas. Como f es por hipótesis regular o semirregular, podemos encontrar un intervalo k -isométrico $I(z)$ que contiene a z y con diámetro positivo. $I(z) \cap P$ debe ser vacío, o, nuevamente por continuidad sería $z \in P$. Así tomando $I(z)$ de diámetro positivo y suficientemente pequeño para que esté incluido en S' , obtenemos el intervalo requerido para que g sea i.l.b. ■

Llegamos así a la conclusión-resumen de este apartado: *Si f es una i.l.b. en S , podemos retirar primero todos los ciclos de los puntos de cambio, despues todos los haces de ciclos cerrados, de forma que la restricción de f al resto de S sea i.l.b. Si además esta restricción es regular (bastaría que f lo fuera en S), podemos retirar todos los ciclos, y la restricción de f al resto del dominio es i.l.b. Llamaremos parte cíclica de f en S al conjunto C unión de todos los ciclos. El conjunto $S-C$ será la parte acíclica de S . Como en la parte cíclica puede haber infinitos ciclos de*

periodos diferentes, no puede afirmarse en general que la restricción de f a S sea periódica, aunque si lo es por puntos. Queda con esto completo el estudio de la restricción de f a su parte cíclica. En los proximos apartados se estudiará la restricción de f a su parte acíclica.

2.4. ISOMETRIAS LOCALES BIYECTIVAS ACICLICAS

En el presente apartado se estudian dos propiedades globales de las i.l.b. acíclicas, que como indica su nombre, son aquellas que no tienen ciclos. El apartado consta de dos teoremas que permiten la reducción de las i.l.b. acíclicas a una clase de aplicaciones de estructura mas sencilla, las reordenaciones por intervalos. Más en particular se probará que podemos retirar los puntos aislados del dominio, como antes hicimos con los puntos cíclicos, y que el número de componentes conexas del resto del dominio es finito. También se pone de manifiesto que las i.l.b. están intimamente relacionadas con una operación topológica elemental, la suma conexa de intervalos. A lo largo de todo el apartado f es una i.l.b. acíclica en S .

2.4.1

Probaremos el siguiente teorema:

i) El número de componentes conexas de S no reducidas a un punto, es finito.

Para demostrarlo, hagamos la hipótesis del absurdo de suponer que tal número es infinito.

Escribamos $S = P \cup V$, donde P agrupa todos los puntos aislados, y V las componentes conexas de diámetro positivo. Destaquemos el conjunto W incluido en V de todas las componentes conexas que contienen algún punto de cambio. Tal conjunto debe ser un conjunto finito, ya que el número de puntos de cambio lo es, y por tanto también el de componentes isométricas (ver IV de 2.2.3). Sea entonces θ el diámetro menor entre los de las componentes isométricas de W y $V(\theta)$ la unión de todas las componentes isométricas cuyo diámetro es mayor o igual a θ .

Para ordenar la demostración del teorema, la dividiremos en etapas:

a) $f(V(\theta)) \subset V(\theta)$

En efecto, $V(\theta)$ se escribe como unión de componentes isométricas. Sea J una de ellas. Como $f(J)$ es un intervalo, debe estar incluido en cierta componente conexa C . Si en C no hay puntos de cambio, entonces C es una componente isométrica y $\theta \leq d(J) = d(f(J)) \leq d(C)$ luego C está en $V(\theta)$, y por tanto también $f(J)$. Si C tiene puntos de cambio, está en W , y se puede poner como unión de componentes

isométricas de diámetro mayor que θ , luego C está en $V(\theta)$, y a fortiori, $f(J)$ también lo está. Por tanto todo $f(V(\theta))$ está incluido en $V(\theta)$.

b) $f(V-V(\theta)) \subset V-V(\theta)$; $L(V-V(\theta)) > 0$ (L = medida Lebesgue)

$V(\theta)$ es unión finita de intervalos, o no sería acotado. Por tanto, es medible. Sabemos por 2.2.1 que $L(V(\theta)) = L(f(V(\theta)))$. Se deduce entonces que $L(V(\theta) - f(V(\theta))) = 0$. Consideremos la siguiente partición de S :

$$S = P \cup (V-V(\theta)) \cup V(\theta)$$

Según la hipótesis del absurdo, como en V hay infinitas componentes conexas de diámetro positivo y en $V(\theta)$ hay sólo un número finito, en $V - V(\theta)$ quedan componentes de diámetro positivo, por lo que dicho conjunto tendrá medida positiva. Queda probar la inclusión. Sea C una componente conexa de $V-V(\theta)$ de diámetro positivo. C es también componente isométrica, ya que en C no puede haber puntos de cambio por estar estos en $W \subset V(\theta)$. Por tanto $f(C)$ es un intervalo que debe estar incluido en una componente conexa M , a su vez incluida en alguno de los tres conjuntos de la partición de S , ya que cada uno de los tres se escribe como unión de componentes conexas de S . M no puede estar en P , pero tampoco en $V(\theta)$, ya que entonces sería $f(C)$ incluido en $V(\theta)$ y como $C \cap V(\theta) = \emptyset$, $f(C) \cap f(V(\theta)) = \emptyset$ y tendría que ser $f(C) \subset V(\theta) - f(V(\theta))$, lo que es contradictorio con el hecho de que la medida de $f(C)$ no pueda ser nula, y la de $V(\theta) - f(V(\theta))$ lo sea. De esta forma se obtiene $f(V-V(\theta)) \subset V-V(\theta)$.

c) En $V-V(\theta)$ hay un punto periódico.

Si conseguimos probar este último resultado, habremos hallado la contradicción deseada con nuestra hipótesis de ser f acíclica, contradicción que rebate nuestra hipótesis del absurdo, y prueba el teorema.

Sea C_1 una componente conexa de $V-V(\theta)$ con diámetro positivo. El número de componentes conexas de $V-V(\theta)$ con diámetro mayor o igual al de C_1 no puede ser infinito, o $V-V(\theta)$ no sería acotado. Como en $V-V(\theta)$ las componentes conexas son también isométricas, $f(C_1)$ será también un intervalo que debe estar incluido en otra componente conexa C_2 con diámetro mayor o igual que C_1 . Reiterando el proceso, obtenemos una cadena C_i de componentes de componentes conexas verificando $f(C_i) \subset C_{i+1}$, $\delta(C_i) \leq \delta(C_{i+1})$. No puede haber infinitas C_i distintas, por lo que para ciertos $m, n \in \mathbb{N}$, $m < n$, $C_m = C_n$. Ello significa, si $k = n-m$, $f^k(C_m) = C_n = C_m$, y aplicando a f^k/C_m el teorema del punto fijo de Brouwer, se obtiene la existencia de un x en C_m tal que $f^k(x) = x$. ■

Observación: De este resultado se desprende que $V = V(\theta)$ y $S = P \cup V(\theta)$, ya que para alcanzar la contradicción hallada basta suponer que $V-V(\theta)$ no es vacío.

2.4.2 El problema de los puntos aislados

Es claro por la definición de i.l.b. que los puntos no aislados de S no pueden tener por imagen puntos aislados. Queremos probar en este punto que la recíproca también es cierta, es decir, el siguiente teorema:

i) La imagen de un punto aislado, ha de ser un punto aislado.

Dos son las consecuencias de este teorema. En primer lugar permite retirar del dominio los puntos aislados, reduciendo el estudio de las i.l.b. al caso en que el dominio consta de un número finito de intervalos de diámetro no nulo. En segundo lugar permitirá probar que la inversa de una i.l.b. es una i.l.b.

Para demostrar el teorema, debemos apoyarnos en la existencia de una invariante topológica de las familias finitas de intervalos respecto a una operación: La suma conexa de intervalos:

Sean J_1 y J_2 intervalos disjuntos. Un intervalo J es suma conexa de J_1 y J_2 y escribimos $J = J_1 \# J_2$ si y solo si J se pone como unión disjunta de dos intervalos I_1 e I_2 , homeomorfos respectivamente a J_1 y J_2 . De esta forma $J_1 \# J_2$ es un intervalo homeomorfo al resultado de "pegar" uno a continuación de otro los intervalos J_1 y J_2 .

La suma conexa de dos intervalos cualesquiera no siempre puede realizarse. Admitiremos como evidente que sólo puede realizarse la suma conexa de dos intervalos "pegando" un extremo cerrado de un intervalo con un extremo abierto del otro. Como la operación está definida salvo homeomorfismo, sólo existen desde el punto de vista de la suma conexa tres tipos de intervalos, el intervalo cerrado que denominaremos c , el abierto que denominaremos a , y el semicerrado que escribiremos s . Las cuatro sumas conexas que pueden hacerse son:

$s \# s = s$
 $s \# a = a$
 $s \# c = c$
 $a \# c = s$

Obsérvese que s actúa a modo de elemento neutro en la operación.

Sea ahora M una familia cualquiera de intervalos disjuntos, en la que hay p intervalos semicerrados, q abiertos y r cerrados. Entonces diremos que M es de tipo (p,q,r) . Pues bien, $\Delta(M) = q - r$ es una invariante de las familias de intervalos respecto de las sumas conexas, en el sentido de que para cualquier familia M' que se puede formar por sumas conexas consecutivas de los intervalos de M , será $\Delta(M') = \Delta(M)$. Para comprobarlo basta ver que cada una de las cuatro sumas conexas indicadas en la tabla de arriba preserva la diferencia entre el número de intervalos abiertos y cerrados de la familia.

Sea ahora f una i.l.b. acíclica en S . Consideremos como en el punto anterior $S = P \cup V$, donde P es el conjunto de los puntos aislados de S , y V agrupa todas las componentes isométricas de diámetro positivo. Sabemos por la observación final del punto anterior que V es unión finita de intervalos disjuntos. Por la definición de i.l.b., $f(V) \subset V$. Como f conserva la medida, $m(V-f(V)) = 0$. Como $f(V)$ se pone también como unión finita de intervalos (de las imágenes de las componentes isométricas de V), $V - f(V)$ será una unión finita de intervalos de medida cero, es decir una unión de un número finito de puntos. Llamemos W a esta familia de puntos. Tendremos dos descomposiciones de V en unión disjunta de intervalos: una como unión de la familia M de las componentes isométricas. Otra como unión de la familia M' formada a su vez por dos colecciones de intervalos, por la colección U de las imágenes por f de las componentes isométricas de V , es decir, de los elementos de M y por la colección de los puntos aislados de la familia W . Así V se obtiene por sumas conexas de las familias M y M' . De aquí deducimos que $\Delta(M) = \Delta(M')$. Como f restringida a los elementos de la familia M es un homeomorfismo es claro que $\Delta(M) = \Delta(U)$ y por tanto $\Delta(M') = \Delta(U)$. Para que esto sea posible, en W debe haber tantos intervalos cerrados como abiertos, pero ya que se trata de una unión finita de puntos, no contiene ningún abierto, y por lo tanto tampoco contiene ningún punto. Esto prueba que W es vacía y que $f(V) = V$. La inyectividad de f da también $f(P) = P$. ■

Es evidente que f/V y f/P son i.l.b. El caso degenerado de f/P no parece interesante desde el punto de vista de la teoría de iteraciones. Pasaremos en lo sucesivo a ocuparnos del caso f/V .

Nota : De la hipótesis del teorema demostrado, se puede suprimir que f sea acíclica. Esta suposición ha sido usada para obtener V como unión finita de intervalos. Al mismo resultado se puede llegar de la siguiente manera: Suprimimos primeramente todos los haces de ciclos cerrados, ya que es evidente que ningún punto aislado puede tener imagen en uno de ellos. Repitiendo ahora el razonamiento del teorema del punto 2.4.1 para la restricción de f al resto, sin suponer que es acíclica, se puede encontrar un haz cíclico cerrado, el correspondiente al elemento periódico que da el apartado c) de la demostración. Este absurdo prueba que en V no puede haber infinitas componentes conexas de diámetro positivo.

2.5. REORDENACIONES POR INTERVALOS.

En este apartado estudiaremos las propiedades globales de las reordenaciones por intervalos, a cuyo estudio nos han conducido los resultados de los anteriores apartados. Esta es su definición:

2.5.1 Definición.

$f: S \rightarrow S$ es una reordenación por intervalos, si es una i.l.b. en S y S es unión finita de intervalos de diámetro positivo. Para las reordenaciones por intervalos escribimos r.i.. Si además son acíclicas, escribiremos r.i.a..

Si la restricción de f a sus intervalos isométricos es creciente, entonces se dice que f es orientable.

La actuación de una r.i. es como sigue: El conjunto finito de sus puntos de discontinuidad, $X(f)$, divide las componentes conexas de S en componentes isométricas, que son transportadas isométricamente por f de forma que recomponen el mismo dominio S . Por tanto, la aplicación f "baraja" S por intervalos, cambiando tal vez el orden de algunos de los mismos.

2.5.2. Inversa de una r.i.

Es inmediato comprobar que la inversa de una r.i. también es una r.i. Las componentes isométricas de f^{-1} son precisamente las imágenes por f de las componentes isométricas de f . Para los puntos de cambio de f^{-1} se obtiene:

$$i) X(f^{-1}) = [f(\text{Fr}(S) \cap S) \cup f(X(f))] \cap \text{Int}(S)$$

Ya que (ver 2.2.2.) los citados puntos de cambio son los extremos de las componentes isométricas de f^{-1} que son interiores a S y se puede probar que dichos puntos son precisamente los del segundo miembro de la igualdad de más arriba.

También se obtienen las siguientes propiedades:

$$ii) f^{-1}(\text{Fr}(S) \cap S) \cap \text{Int}(S) \subset X(f)$$

En efecto, si la imagen por f de un punto x interior a S es un punto frontera de S y x no fuera punto de cambio, podríamos encontrar un entorno abierto de x que fuera un intervalo isométrico (ver 2.2.2.), y esto es incompatible con que $f(x)$ esté en la frontera de S . Si además f es acíclica se obtiene:

$$iii) \text{ Si } x \in \text{Fr}(S) \cap S, \text{ para algún } n \in \mathbb{N}, f^{-n}(x) \in X(f)$$

Ya que como $\text{Fr}(S) \cap S$ es un conjunto finito las $f^{-1}(x)$ no pueden pertenecer todas a este conjunto habría ciclos. Si $y = f^{-k}(x)$ es la primera que no está en la frontera de S , aplicando ii) a $z = f^{-1}(y)$ se obtiene el resultado.

Si además f es regular o semirregular (ver definiciones en 2.3.6) se obtiene:

iv) Las iteraciones de f son r.i. Más en general, las aplicaciones f^k , $k \in \mathbb{Z}$, son r.i. y forman un grupo con respecto a la composición de funciones, que llamaremos grupo iterativo generado por f . Si f es acíclica, también lo son todas las r.i. de su grupo iterativo. Para los puntos de cambio se obtiene:

$$X(f^k) \subset \left(\bigcup_{i=0}^{k-1} f^{-i}(X(f)) \right) \cap \text{Int}(S)$$

$$X(f^{-k}) \subset \left[\bigcup_{i=1}^k f^i(X(f)) \cup f^k(\text{Fr}(S)) \cap S \right] \cap \text{Int}(S)$$

Finalmente, si f y g son r.i. en S , se obtiene la siguiente condición para que su composición lo sea:

v) $g \circ f$ es r.i. si $X(f^{-1}) \cap X(g) = \emptyset$. También lo es si se verifica $f^n(X(f)) \cap Xg = \emptyset$, $n \in \mathbb{N}$. Para los puntos de cambio de $g \circ f$ se obtiene:

$$X(g \circ f) \subset (f^{-1}(X(g)) \cup X(f)).$$

2.6. PROPIEDADES DE LAS REORDENACIONES POR INTERVALOS ACICLICAS Y SEMIRREGULARES.

2.6.1. Operadores ω y α asociados a f

Sea f una r.i. en S . Definimos la órbita $O(x)$ de un elemento x de S en la misma forma que en 2.1.1., tomando también idéntica definición para $O^+(x)$ y $O^-(x)$. Definimos también el conjunto $\omega(x)$ como el conjunto de todos los valores de adherencia de $O^+(x)$. Se puede definir ω como operador de conjuntos. Dado $A \subset S$ definimos:

$$\omega(A) = \bigcap_{i=0}^{\infty} \left(\bigcup_{n=i}^{\infty} f^n(A) \right),$$

Donde la adherencia es relativa a S .

Análogamente se define el operador α :

$$\alpha(A) = \bigcap_{i=0}^{\infty} \left(\bigcup_{n=i}^{\infty} f^{-n}(A) \right)$$

Si fuera necesario aclarar la aplicación cuyas iteraciones intervienen en la definición de estos operadores, escribiríamos $\omega_f(A)$, $\omega_g(A)$ etc.

Decimos que f es transitiva, si para algún $x \in S$ es $\omega(x) = S$, y decimos que S es minimal si $\forall x \in S$ es $\omega(x) = S$.

Indicamos a continuación algunas propiedades de los operadores α y ω que se desprenden directamente de las definiciones de los mismos.

$$i) \omega_f = \alpha_f^{-1}; \alpha_f = \omega_f^{-1}$$

$$ii) \omega(A) \text{ y } \alpha(A) \text{ son } S\text{-cerrados.}$$

Estas propiedades serían ciertas si f fuera cualquier biyección en un espacio topológico S . En la próxima se requeriría que en S hubiera bases numerables de entornos. Por ejemplo, que fuera un espacio métrico.

$$iii) \omega(A \cup B) = \omega(A) \cup \omega(B)$$

$$\alpha(A \cup B) = \alpha(A) \cup \alpha(B).$$

Para obtener esta propiedad, se necesita el siguiente lema, útil en muchas ocasiones:

IV) $x \in \omega(A) \Leftrightarrow$ para todo entorno $V(x)$ de x , podemos encontrar una sucesión $z = \{z_n\} \subset A$ verificando que para cualquier natural n , hay otro natural k mayor que n tal que $f^k(z_n) \in V(x)$. Este lema se deriva directamente de la definición de ω . Otra forma de expresar el mismo resultado, es diciendo que en cada $V(x)$ hay infinitos términos de la forma $f^k(z_n)$, $k > n$. Si A se reduce a un punto x , se requiere simplemente que haya infinitos términos de la forma $f^k(x)$. Para el operador α se obtienen condiciones análogas.

Demostraremos ahora la propiedad iii) para el operador ω . Para el operador α la demostración es análoga. La inclusión \subset es inmediata, ya que:

$$\begin{array}{l} \omega(A) \subset \omega(A \cup B) \\ \omega(B) \subset \omega(A \cup B) \end{array} \quad \Rightarrow \quad \omega(A) \cup \omega(B) \subset \omega(A \cup B)$$

Para la demostración de la inclusión contraria, tomamos $x \in \omega(A \cup B)$, y consideramos la familia $\{B(k)\}$ de bolas centradas en x y de radio $1/k$. Según el lema IV) para cada $B(k)$ podemos encontrar una sucesión $z(k) = \{z_n(k)\}$ de términos de $A \cup B$ con infinitos términos de la forma $f^i(z_n(k))$ con $i > n$ en $B(k)$. O bien podemos formar una subsucesión de la $z(k)$ que tiene todos sus términos en A o bien podemos formar una subsucesión con todos sus términos en B , de forma que podemos suponer que cada $z(k)$ consta exclusivamente de términos de A o de B . O bien hay infinitas $z(k)$ que tienen todos sus términos de A , o bien hay infinitas que tienen todos sus términos de B . En el primero de los casos, si dicha colección de valores

de K es la sucesión $\{K_j\}$, la sucesión de términos de A , $\{y_j\}$, de término general $y_j = z_j(K_j)$, tiene la propiedad de que hay infinitos términos de la forma $f^i(y_j)$, $i > j$ en cada $B(K)$ y por tanto en cada entorno de $V(x)$ del punto x , y de acuerdo con el lema IV), $x \in \omega(A)$. En el segundo caso, se obtiene igualmente $x \in \omega(B)$.

$$V) \omega(A \cap B) \subset \omega(A) \cap \omega(B); \alpha(A \cap B) \subset \alpha(A) \cap \alpha(B)$$

Ambas son inmediatas, pero no así las inclusiones inversas, que no son ciertas, ya que más adelante encontraremos que para elementos $x \neq y$ de S puede ser $\omega(x) = \omega(y)$.

$$VI) \text{ Si } n \in \mathbb{N} \text{ es } \omega(A) = \omega(f^n(A))$$

Ya que de las sucesiones requeridas en la propiedad IV) sólo se suprimen en el segundo miembro de VI) un número finito de términos.

2.6.2. Propiedades de α y ω asociadas a las propiedades de una r.i.

En el punto anterior se han visto propiedades de los operadores α y ω generales, comunes a cualquier biyección en un espacio topológico - en el caso de iii) en un espacio métrico -, pero no se han utilizado en ningún momento las propiedades de una r.i. En este punto se estudian algunas propiedades de dichos operadores que sí dependen de las propiedades de las r.i. En todo el punto suponemos que f es una r.i. en S .

$$i) \text{ Si } x \in S - U f^{-n}(X(f)), x \in \omega(y), \text{ entonces } \omega(x) \subset \omega(y).$$

La razón de esta propiedad es que las f^k son continuas en x para cualquier $k \in \mathbb{N}$. En cualquier entorno $V(z)$ de un punto z de $\omega(x)$, hay infinitos términos de la forma $f^k(x)$. Como x está en $\omega(y)$, cerca de x hay infinitos puntos de la forma $f^i(y)$, y por la continuidad de f^k en x , también los hay en $V(z)$, y por tanto z está en $\omega(y)$. ■

ii) Si f es acíclica, sus intervalos ω -isométricos (ver 2.3.3.) se reducen a puntos.

En efecto, consideremos la existencia de un intervalo ω -isométrico H de diámetro positivo. Podemos elegir H maximal, en el sentido de que no está estrictamente incluido en ningún intervalo ω -isométrico. Consideramos la familia de intervalos ω -isométricos $\{H_k\}$, $k \in \mathbb{N}$, donde $H_k = f^k(H)$. Además todos ellos deben ser maximales, o de lo contrario, por la continuidad de las restricciones de las f^{-k} a los intervalos ω -isométricos, se obtendría que H no es maximal. Ahora bien ello significa que H_k y $H_{k'}$, si k y k' son distintos, o no se cortan o coinciden, ya que si se cortaran sin coincidir, su unión sería un intervalo ω -isométrico que los contendría, en contra de su maximalidad. No es posible que los H_k sean disjuntos dos a dos, ya que todos ellos tienen un mismo diámetro positivo, y S sería no acotado. Pero si para ciertos k, k' , $k > k'$, fuese $H_k = H_{k'}$, para $p = k - k'$ se tendría $f^p(H_{k'}) = H_{k'}$, y siendo

$f|_{P/H_K}$ una isometría, aplicando el teorema de Brouwer se obtendría la existencia en H_K de un punto cíclico, contra la hipótesis de que f es acíclica. ■

iii) Si f es acíclica, $\alpha(X(f)) = S$

Para probarlo, tomemos un punto x del conjunto $Y = S - \alpha(X(f))$, suponiendo que Y no es vacío. Por la definición de α , podemos encontrar un S -entorno de x , $V(x)$, tal que para cierto n :

$$V(x) \cap \left(\bigcup_{i=n}^{\infty} f^{-i}(X(f)) \right) = \emptyset$$

Supondremos además que $V(x)$ es S -abierto. En estas condiciones el conjunto:

$$V(x) = \bigcup_{i=0}^{n-1} f^{-i}(X(f))$$

Es S -abierto y no contiene puntos de cambio para ninguna f^k , $k \in \mathbb{N}$, de forma que cualquiera de sus componentes conexas, es un intervalo w -isométrico de diámetro positivo, contra la propiedad ii). ■

Análogamente, se obtiene para el operador w :

iv) Si f es acíclica, $w(X(f)) = S$

Esta propiedad es consecuencia de la anterior y de las relaciones entre los operadores α y w . Para probarlo, empecemos recordando que según la propiedad i) de 2.5.2. los puntos de cambio de f^{-1} están incluidos en el conjunto $Y = f(\text{Fr}(S) \cap S) \cup f(X(f))$. Probaremos primero que $S \subset w(Y)$. En efecto, aplicando la propiedad anterior al operador α_f^{-1} se obtiene:

$$\alpha_f^{-1}(X(f^{-1})) = S.$$

Pero como $X(f^{-1}) \subset Y$, obtenemos:

$$S \subset \alpha_f^{-1}(Y) = w_f(Y)$$

Donde para la última igualdad se ha utilizado la propiedad i) de 2.6.1.

Probaremos a continuación $w(Y) \subset w(X(f))$, y ello nos dará el resultado final deseado, ya que después de lo anteriormente probado, resultará que $S \subset w(X(f))$, y la inclusión contraria es obvia. Por la propiedad iii) de los operadores α y w (2.6.1.):

$$w(Y) = \left[\bigcup_{x \in X(f)} w(f(x)) \right] \cup \left[\bigcup_{x \in \text{Fr}(S) \cap S} w(f(x)) \right]$$

En cuanto al conjunto del primer corchete, por la propiedad Vi) del mismo apartado, para cada $x \in X(f)$ es $w(f(x)) = w(x)$, de forma que:

$$\bigcup_{x \in X(f)} \omega(f(x)) = \bigcup_{x \in X(f)} \omega(x) = \omega(X(f))$$

En cuanto al conjunto del segundo corchete, aplicando el mismo razonamiento:

$$\bigcup_{x \in \text{Fr}(S) \cap S} \omega(f(x)) = \bigcup_{x \in \text{Fr}(S) \cap S} \omega(x)$$

Pero utilizando la propiedad iii) de 3.5.2. sabemos que si $x \in \text{Fr}(S) \cap S$, podemos encontrar $y \in X(f)$ tal que $f^n(y) = x$ para cierto natural n , de forma que $\omega(x) = \omega(f^n(y)) = \omega(y) \subset \omega(X(f))$, luego:

$$\bigcup_{x \in \text{Fr}(S) \cap S} \omega(f(x)) \subset \omega(X(f)).$$

Y esto concluye la prueba de que $\omega(Y) \subset \omega(X)$. ■

2.6.3. Teorema de recurrencia.

El resultado que acabamos de exponer, es algo mas débil que el de la transitividad: No podemos afirmar que para cierto elemento x de S el ω -conjunto de x sea todo S . Pero sí podemos afirmar que al menos sucede que el ω -conjunto de un conjunto finito, $X(f)$, coincide con S . Es fácil por otra parte obtener ejemplos de r.i.a. que no es minimal ni transitiva. Ejemplos de este tipo se pueden también obtener para el caso particular de los intercambios de intervalos. Para obtener resultados más fuertes, es necesario introducir una nueva hipótesis para f : la regularidad o semirregularidad. En este punto, introduciendo la hipótesis más débil de semiregularidad (ver 2.3.6.) se obtiene uno de los teoremas básicos del capítulo, que llamaremos teorema de recurrencia, porque constituye un refinamiento del teorema de recurrencia de Poincaré aplicado al caso que nos ocupa. Dicho teorema establece que si $T: X \rightarrow X$ es una aplicación Borel medible en un espacio métrico separable X y μ es una medida T -invariante en la σ -álgebra de Borel de X , entonces μ -casi todos los puntos de X son recurrentes, es decir:

$$\mu \{ x : x \text{ no pertenece a } \omega_T(x) \} = 0$$

Ahora bien, si f es una r.i.a. en S , todas las hipótesis son satisfechas, tomando como medida m la medida de Lebesgue. Por consiguiente, el conjunto de puntos no recurrentes en una r.i.a. debe tener medida nula. Por su parte el teorema de recurrencia para una r.i.a. semiregular establece que:

i) Si f es una r.i.a. semiregular en S , el conjunto de puntos no recurrentes es vacío, más aún, $\forall p \in \mathbb{N} \cup \{0\}$, $f^p(x) \in \omega(x)$.

Para probarlo, supongamos que hay un $z = f^p(x)$, y que z no pertenece a $\omega(x)$.

En esta hipótesis podemos encontrar un intervalo S -abierto tal que

$z \in V(z)$, que sólo contiene un número finito de puntos de la forma $f^k(x)$, $k \in \mathbb{N}$ (ver iv de 2.6.1.). Reduciendo convenientemente $V(z)$ podemos conseguir que en él no haya sino un sólo punto de dicha forma: el mismo z . Además como f es acíclica si $p \neq k$, $f^p(x) \neq f^k(x)$.

Sabemos por otra parte por la hipótesis de semirregularidad que hacemos en f (2.3.6.) $\forall k \in \mathbb{N}$, podemos encontrar un intervalo J_k k -isométrico y con diámetro positivo. Lo podemos hacer de forma que sea $J_k \subset J_{k'}$ si $k > k'$. Consideremos la sucesión de intervalos $\{H_k\}$ donde $H_k = V(z) \cap J_k$. Es una sucesión de intervalos de diámetro positivo. Además como cada H_k es k -isométrico, los H_k se contraen a z cuando $n \rightarrow \infty$, o de lo contrario obtendríamos un intervalo w -isométrico de diámetro positivo, contra ii) de 2.6.2.

Pongamos ahora $G_m = \text{Int}(H_m) \cup \{z\}$ y definamos:

$$i_m = \min \{ i \in \mathbb{N} : f^{-i}(X(f)) \cap G_m \neq \emptyset \}$$

Tal mínimo existe, o G_m sería un intervalo w -isométrico.

Tomemos para cada m un $y_m \in f^{-i_m}(X(f)) \cap G_m$. Tal y_m verifica:

$$f^{i_m}(y_m) \in X(f).$$

Además por la condición de mínimo:

$$\bigcup_{i=0}^{i_m-1} f^{-i}(X(f)) \cap G_m = \emptyset$$

Por lo que f^{i_m}/G_m es una isometría.

Como $X(f)$ es un conjunto finito, para infinitos i_m se debe repetir algún valor de $f^{i_m}(y_m)$. Sea éste a y consideremos la subsucesión $\{m_k\}$ de valores de m para los que $f^{i_m}(y_m) = a$. Definimos las sucesiones $\{J_k\}$, $\{t_k\}$, $\{I_k\}$ de términos generales:

$$J_k = i_{m_k}; t_k = y_{m_k}; I_k = G_{m_k}$$

Sucesiones para las que se verifican las relaciones:

$$t_k \longrightarrow z$$

$$f^{J_k}(t_k) = a$$

$$f^{J_k}/I_k \text{ es isometría}$$

Para un valor de k lo suficientemente grande, $f^{J_k}(z) \neq a$, ya que en caso contrario a sería cíclico. Esto prueba que si k es mayor que cierto K , los t_k son diferentes de z y además están en el interior de los I_k (ver definición de estos intervalos y de los G_k).

Sea $n > K$. Como $t_n \in \text{Int}(I_n)$ y $t_n \neq z$, podemos encontrar una bola $B(t_n) \subset I_n$ centrada en t_n en la que no está z ni hay términos de la forma $f^k(x)$, $k > 0$ (z podría ser el único elemento de esta forma por la definición de los intervalos H_m). Resulta entonces que $f^{j_n}(B(t_n)) = B(a)$ es una bola centrada en a en la que no debe haber elementos de la forma $f^k(x)$ para $k > j_n$, ya que en caso contrario habría elementos de la forma $f^k(x)$, con $k > 0$ en $B(t_n) = f^{-j_n}(B(a))$. De aquí surge la contradicción, porque tomando una bola $B(z)$ de centro en z y de radio igual al de $B(a)$, podemos seleccionar un natural q suficientemente grande para que $I_q \subset B(z)$ y para que $j_q > j_n$. Como f^{j_q} actúa isométricamente sobre I_q y la distancia de z a t_q es menor al radio de $B(a)$, $f^{j_q}(z)$ debe estar en esta bola, y como $z = f^p(x)$, resulta que tomando $k = p + j_q > j_n$, encontramos en $B(a)$ un elemento de la forma $f^k(x)$, $k > j_n$. ■

A partir del teorema de recurrencia se obtiene el teorema análogo para el operador α :

ii) Si f es una r.i.a. semirregular en S y $x \in S$, entonces $\forall p \in \mathbb{N}$, $f^{-p}(x) \in \alpha(x)$.

Basta para probarlo aplicar el teorema de recurrencia al operador ω_f^{-1} . También se obtiene inmediatamente el siguiente corolario, que simplifica, en la hipótesis de ser f una r.i.a. semirregular, la definición de los operadores α y ω :

iii) $\omega(x)$ coincide con la S -adherencia de $O^+(x) \cup \{x\}$; $\alpha(x)$ coincide con la S -adherencia de $O^-(x)$.

2.6.4. Teorema de estructura de $\omega(x)$ en una r.i.a. regular.

En este apartado se obtendrá el resultado fundamental respecto a la estructura topológica de los conjuntos $\omega(x)$ y $\alpha(x)$ de cualquier $x \in S$ si f es una r.i.a. semirregular en S . Después del anterior apartado, en concreto del corolario iii), ello es equivalente a la caracterización topológica de las órbitas de los elementos de S . El teorema de estructura establece que:

i) Si f es una r.i.a. semirregular en S , $\forall x \in S$ $\omega(x)$ - y por consiguiente también $\alpha(x)$ - se pone como unión de intervalos S -cerrados de diámetro positivo.

Previamente a la obtención de este resultado se precisa el siguiente resultado, de fácil demostración.

ii) Si f es una r.i. en S y $z \in S$ verifica: $z \in \omega(x) - X(f)$, entonces $f(z) \in \omega(x)$.

Basta tener en cuenta las definiciones y el hecho de que f es continua en z .

Pasamos ahora a la demostración del teorema de estructura. La demostración se basa en el análisis de la estructura del conjunto $V = S - \omega(x)$. Para clarificar la demostración, la dividimos por

etapas. Probaremos primero :

a) *V se escribe como unión numerable de sus componentes conexas. Estas son S-abiertas, y tienen su S-frontera sobre $\omega(x)$.*

El conjunto V es S-abierto por ser $\omega(x)$ S-cerrado. Esto prueba que puede ponerse como unión numerable de S-abiertos, que convenientemente agrupados dan las componentes conexas S-abiertas de V. Sea J una de ellas. Será:

$$\text{Fr}(J) \cap S = \overline{J \cap S \cap \omega(x)}$$

Y como $\omega(x)$ es S-cerrado, $\overline{\omega(x) \cap S} = \omega(x)$, de forma que:

$$\text{Fr}(J) \cap S \subset \omega(x).$$

b) *El número de componentes conexas de V es finito.*

Sabemos que $V = \bigcup_{i \in M} J_i$, donde M es un conjunto numerable o finito y

y las J_i las componentes conexas de V. Como el conjunto $Y = \text{Fr}(S) \cup U \cap X(f)$ es finito, sólo un número finito de componentes J_i pueden tener en su adherencia puntos de Y. Sea θ el menor de los diámetros de estas J_i , necesariamente positivo. Consideramos el conjunto de índices $P \subset M$ tal que $i \in P \Rightarrow d(J_i) \geq \theta$. Es claro que P es finito o no sería S-acotado. Para probar, como deseamos, que M es finito es suficiente probar que $Q = M - P$ es vacío. Probaremos que esto es cierto por reducción al absurdo.

Sea pues $i \in Q$. Es claro por las definiciones que $J_i \cap Y = \emptyset$. De:

$$\overline{J_i \cap X(f)} = \emptyset$$

Se deduce que f/J_i es isometría. Probaremos que $\text{Fr}(J_i) \subset \omega(x)$. En efecto, los extremos de las J_i son puntos frontera de S, o bien pertenecen a la frontera relativa de J_i en S. El primer supuesto es imposible pues sabemos que :

$$\overline{J_i \cap \text{Fr}(S)} = \emptyset$$

Esto significa que $\text{Fr}(J_i) \subset \text{Fr}(S) \cap S \subset \omega(x)$, según a).

Por otra parte los extremos de J_i no pueden ser puntos de cambio, por lo que se verifica por la continuidad de f en dichos puntos :

$$f(\text{Fr}(J_i)) = \text{Fr}(f(J_i))$$

Como $\text{Fr}(J_i) \subset \omega(x) - X(f)$, resulta por lo anterior e ii) que también $\text{Fr}(f(J_i)) \subset \omega(x)$. Esto prueba que $f(J_i)$ es un intervalo que tiene sus extremos en $\omega(x)$.

Por otra parte $f(J_i) \subset V$ ya que alrededor de cada $y \in J_i$ podemos encontrar una S -bola $B(y)$ sin puntos de la forma $f^k(x)$ y que se transforma isométricamente por f , lo que prueba que en $f(B(y))$, que es una S -bola alrededor de $f(y)$, no puede haber infinitos puntos de la forma $f^k(x)$, y por tanto y no puede estar en $w(x)$.

Se concluye que $f(J_i)$ es un intervalo incluido en V y con extremos en $w(x)$, de donde no puede estar incluido en ningún otro intervalo de V . Se trata por tanto de un intervalo maximal de V , es decir, de una componente conexa J_j . El diámetro de J_j es igual al de J_i , y por tanto menor que θ , de donde concluimos que $j \in Q$.

Reiterando el mismo razonamiento, concluimos que todos los $f^k(J_i)$, $k \in \mathbb{N}$ son de la forma J_n con $n \in Q$. Pero esto es imposible, porque sería J_i un intervalo w -isométrico, contra lo probado en ii) de 3.6.2. De esta contradicción se deduce que $Q = \emptyset$, y por tanto M es finito, y V es unión finita de intervalos S -abiertos.

c) $\text{Int}_S(w(x)) \neq \emptyset$

Pues de lo contrario $w(x)$ sería un conjunto frontera, y siendo S -cerrado, coincidiría con su S -frontera, igual a su vez a la S -frontera de su complementario. Pero esto es imposible, ya que V , como unión finita de intervalos, tiene una S -frontera finita, y sin embargo, $w(x)$ es un conjunto infinito, ya que por el teorema de recurrencia sabemos que consta al menos de todos los $f^k(x)$, distintos entre sí por ser f acíclica.

d) $w(x)$ es unión finita de intervalos S -cerrados de diámetro positivo.

Que es unión finita de intervalos, se desprende del hecho de que es el complementario de V , unión finita de intervalos, respecto de S , también unión finita de intervalos. Para probar que son de diámetro positivo, basta ver que $w(x)$ no puede tener puntos aislados. Ello se deduce de que en todo S -entorno de $y \in w(x)$ debe haber infinitos puntos de la forma $f^k(x)$, y todos ellos pertenecen, según el teorema de recurrencia, a $w(x)$. Finalmente los intervalos que componen $w(x)$ son claramente S -cerrados. ■

2.6.5. Teorema de minimalidad

En este punto se completa el estudio de las r.i.a. obteniendo una condición necesaria y suficiente de minimalidad para las mismas en el caso de que sean regulares. Probaremos en primer lugar que para dos puntos de cambio x e y , sus omega-conjuntos son esencialmente iguales, o esencialmente diferentes, en concreto:

i) Si f es una r.i.a regular en S , $x \in X(f)$, $y \in S$, entonces $\text{Int}_S(w(x)) \cap \text{Int}_S(w(y)) \neq \emptyset \Rightarrow \text{Int}_S(w(x)) \subset \text{Int}_S(w(y))$. Si además suponemos que también $y \in X(f)$, entonces la anterior inclusión es una igualdad.

Esta segunda afirmación se desprende de la primera, intercambiando los papeles de x e y .

Partimos por tanto de la hipótesis $x \in X(f)$, $y \in S$. Tomaremos $z \in \text{Int}_S(\omega(x))$ y probaremos que $z \in \omega(y)$. De esta forma tendremos que $\text{Int}_S(\omega(x)) \subset \omega(y) \Rightarrow \text{Int}_S(\omega(x)) \subset \text{Int}_S(\omega(y))$, que es lo que queremos concluir.

Sea por tanto $x \in X(f)$, $y \in S$, $z \in \text{Int}_S(\omega(x))$.

Para probar como deseamos que $z \in \omega(y)$, basta probar que dada una S -bola $B_\epsilon(z)$ centrada en z y de radio ϵ arbitrario, y un natural i arbitrario, podemos encontrar un $k > i$ tal que $f^k(y) \in B_\epsilon(z)$.

Tomamos para ello un abierto no vacío $W \subset \omega(x) \cap \omega(y)$, cuya existencia garantizan nuestras hipótesis. Tomamos $n > i$ tal que $f^n(x) = p \in W$. Para ningún $s \in \mathbb{N}$ podrá ser f^s discontinua en p , o de lo contrario f no sería regular- recuerdése que $x \in X(f)$ -.

Como $z \in \omega(x)$, podemos encontrar un $m > n$ tal que $f^m(x) \in B_\epsilon(z)$. Si es $s = m - n$, entonces $f^s(p) = f^m(x) \in B_\epsilon(z)$. Tomemos una S -bola $B_\delta(p)$ con δ lo suficientemente pequeño para que f^s sea isometría en $B_\delta(p)$, y para que $f^s(B_\delta(p)) \subset B_\epsilon(z)$, lo que podemos conseguir ya que sabemos que $f^s(p) = f^m(x) \in B_\epsilon(z)$. Además, como $p \in \omega(y)$, podemos encontrar $t > i$ tal que $f^t(y) \in B_\delta(p)$, con lo que si $k = s + t$, resulta que $f^k(y) = f^s(f^t(y)) \in f^s(B_\delta(p)) \subset B_\epsilon(z)$. ■

Precisamos a continuación de la siguiente proposición auxiliar:

ii) Si f es una r.i.a regular en S , entonces:

$$x \in \bigcup_{i=0}^{\infty} f^{-i}(X(f)) \Rightarrow f^{-k} \text{ es continua en } x \quad \forall k \in \mathbb{N}$$

Es consecuencia de las definiciones y de las relaciones (iii) y (iv) de 2.5.2.

A continuación se obtiene un refinamiento del resultado de i) para el caso en que f sea regular, consiguiendo una caracterización precisa de las intersecciones de los omega-conjuntos de los puntos de cambio.

iii) Si f es una r.i.a. regular en S , y x e y son puntos de $X(f)$, entonces o bien $\omega(x) \cap \omega(y) \neq \emptyset$ o bien $\omega(x) = \omega(y)$.

Probaremos en primer lugar que $\omega(x) \cap \omega(y) \neq \emptyset \Rightarrow$

bien : $\omega(x) = \omega(y)$, o bien :

$$\omega(x) \cap \omega(y) \subset \bigcup_{i=0}^{\infty} f^{-i}(X(f))$$

Supongamos para demostrarlo que la última relación no es cierta, y sea $z \in \omega(x) \cap \omega(y)$. Según i) de 3.6.2. $\omega(z) \subset \omega(x)$ y $\omega(z) \subset \omega(y)$.

luego $\omega(z) \subset \omega(x) \cap \omega(y)$. El teorema de estructura prueba que $\text{Int}_S \omega(z)$ no es vacío, y como $\text{Int}_S(\omega(z)) \subset \text{Int}_S(\omega(x)) \cap \text{Int}_S(\omega(y))$, según la proposición que se acaba de probar se obtiene $\text{Int}_S(\omega(x)) = \text{Int}_S(\omega(y))$. Pero siendo $\omega(x)$ y $\omega(y)$ uniones finitas de intervalos S -cerrados de diámetro positivo, la coincidencia de sus S -interiores es equivalente a la coincidencia de todo el conjunto.

A continuación mostraremos que de la alternativa señalada, sólo es posible la segunda opción. Supongamos para ello que $\omega(x) \cap \omega(y) \neq \emptyset$ y que $\omega(x) \neq \omega(y)$. Según lo probado:

$$\omega(x) \cap \omega(y) \subset \bigcup_{i=0}^{\infty} f^{-i}(X(f))$$

Tomemos $z \in \omega(x) \cap \omega(y)$, $z = f^{-i}(t)$, $i \in \mathbb{N}$, $t \in X(f)$. El punto z debe pertenecer a la frontera común de una componente conexa $C(x)$ de las que constituyen $\omega(x)$ y de otra componente conexa $C(y)$ de las que componen $\omega(y)$, yaciendo cada una a un lado de z , ya que si z estuviera en el S -interior de alguna componente conexa de $\omega(x)$ o de $\omega(y)$, de ello se deduciría que los S -interiores de ambos conjuntos tendrían intersección no vacía, situación incompatible con el supuesto que hacemos sobre $\omega(x) \cap \omega(y)$. Este mismo supuesto, junto con ii) significa la continuidad de f^{-k} $\forall k \in \mathbb{N}$, por lo que $f^{-k}(z)$ pertenece siempre a la S -frontera de $\omega(x)$ y de $\omega(y)$. Pero esto significa que estas S -fronteras son conjuntos infinitos, en contradicción con el teorema de estructura, o bien que f es cíclica, también contradictorio con nuestras hipótesis. ■

El próximo paso consiste en establecer la relación existente entre el ω -conjunto de los puntos de cambio y la componentes conexas de S mediante la siguiente proposición:

iv) Si f es una r.i.a. regular en S , $\forall x \in X(f)$, el conjunto $\omega(x)$ es S -abierto, y se pone como unión de componentes conexas de S . Más concretamente, se pone como unión de aquellas componentes conexas que son "visitadas" por las sucesivas $f^k(x)$, $k \geq 0$.

El resultado anterior junto con iv) de 2.6.2. prueban que S se puede escribir como la unión disjunta de ω -conjuntos de puntos de cambio. Esto prueba que $\omega(x)$, $x \in X(f)$, se obtiene como complementario de una unión finita de ω -conjuntos, y como cada uno de ellos es S -cerrado resulta que $\omega(x)$ es S -abierto. Como cada una de las componentes conexas C de S es S -abierto, su intersección con $\omega(x)$ es C -abierto y C -cerrada, de donde debe ser vacía o coincidir con C . Esto prueba que $\omega(x)$ se escribe como unión de componentes conexas de S . Si en una componente conexa de S no hay ninguna $f^k(x)$, entonces es claro que dicha componente no puede estar incluida en $\omega(x)$. Si, por el contrario, alguna $f^k(x)$ está en una componente conexa de S , por el teorema de recurrencia resulta que la intersección de dicha componente y $\omega(x)$ no es vacía ya que consta al menos de la $f^k(x)$, y por lo tanto toda la componente conexa debe estar incluida en $\omega(x)$, así que $\omega(x)$ se escribe como unión de las componentes conexas de S visitadas por las $f^k(x)$.

$k \geq 0$. ■

Solo resta extender los resultados obtenidos en iii) y iv) a puntos que no son de cambio. Esto se logra como consecuencia directa de la siguiente proposición:

v) Si f es una r.i.a. regular en S , $\forall x \in S$, $\exists t \in X(f)$: $\omega(x) = \omega(t)$. Por tanto lo dicho en iii) y iv) tiene validez para todos los puntos de S .

Primero estudiaremos el caso en que x no pertenece a $\bigcup_{i=0}^{\infty} f^{-i}(X(f))$, y en que, por tanto, f^i es continua en x $\forall i \in \mathbb{N}$.

Como S se obtiene como unión de omega-conjuntos de puntos de cambio, para cierto $y \in X(f)$ es $x \in \omega(y)$, de donde por i) de 2.6.2 resulta que $\omega(x) \subset \omega(y)$. Pero por el teorema de estructura sabemos que $\text{Int}_S(\omega(x)) \neq \emptyset$, y tomando $z \in \text{Int}_S(\omega(x))$ y un $t \in X(f)$ tal que sea $z \in \omega(t)$, será, por ser $\omega(t)$ S -abierto, $z \in \text{Int}_S(\omega(t))$. De aquí que $z \in \text{Int}_S(\omega(x)) \cap \text{Int}_S(\omega(t)) \neq \emptyset$. Aplicando ahora i) de este mismo punto se deduce que $\omega(t) = \text{Int}_S(\omega(t)) \subset \text{Int}_S(\omega(x)) \subset \omega(x) \subset \omega(y)$. De aquí se concluye que $\omega(t) \cap \omega(y) \neq \emptyset$, y por ser t e y puntos de cambio, por iii), que $\omega(t)$ y $\omega(y)$ coinciden entre sí y también coinciden con $\omega(x)$, comprendido entre ambos.

En el segundo caso, si $x \in \bigcup_{i=0}^{\infty} f^{-i}(X(f))$, entonces para cierto $t \in X(f)$ es $f^i(x) = t$, de donde por vi) de 2.6.1 resulta $\omega(x) = \omega(t)$. ■

Con todos estos resultados estamos en condiciones de obtener fácilmente el Teorema de minimalidad para una r.i.a. regular:

vi) Si f es una r.i.a. regular en S , las siguientes afirmaciones son equivalentes :

- a) f es minimal
- b) f es transitiva
- c) $\exists x \in S$ tal que alguna $f^k(x)$, $k \geq 0$, está en cada componente conexa de S .

Si f no es minimal, su dominio se descompone en unión de como máximo tantos conjuntos como componentes conexas tiene S , tales que la restricción de f a cada una de ellos es una r.i. minimal.

a) \Rightarrow b) y b) \Rightarrow c) son inmediatas. En las hipótesis de c), de iv) y v) se desprende directamente b). Finalmente de iii) y v) se deduce que b) \Rightarrow a). La segunda afirmación se desprende del hecho de que la restricción de f a la unión de componentes conexas visitadas por un punto es una r.i.a. minimal, según los resultados anteriores. ■

3.7. GENERALIZACIONES.

En este apartado se obtienen diversas generalizaciones que permitirán posteriormente mayor flexibilidad en las aplicaciones.

2.7.1. Generalización de la definición de r.i.

En las aplicaciones, resulta a veces molesta la exigencia de que todos los intervalos isométricos de una r.i. deban ser de diámetro positivo, y puede ser muy conveniente considerar la posibilidad de que haya un número finito de puntos para los que dicho intervalo se reduzca a los mismos puntos. Este caso se produce, como veremos en el próximo capítulo, en los billares poligonales de ángulos racionales, en donde el rebote en los vértices plantea problemas de esta clase.

Consideremos el siguiente sistema dinámico: Sea S unión finita de intervalos de diámetro positivo y $f:S \rightarrow S$ una aplicación con un número finito de puntos de discontinuidad, cuya restricción al dominio de continuidad es inyectiva y preserva la medida Lebesgue. Esta última hipótesis es equivalente a la actuación isométrica sobre los intervalos de continuidad, que serán así intervalos isométricos. La diferencia con una i.l.b., es que ahora hay un número finito de puntos en los que no se exige continuidad lateral, y sólo se necesita la inyectividad en el complementario de estos puntos. Bajo estas condiciones decimos que f es una reordenación por intervalos generalizada.

Nuestro objetivo es probar que puede construirse una reordenación por intervalos g a cuya dinámica está subordinada la de f .

En una primera etapa probaremos que podemos construir una aplicación $g : S \rightarrow S$ que coincide con f salvo en los puntos de discontinuidad, y tal que es biyectiva.

Consideremos para ello la familia $\{I_i\}$ de intervalos isométricos de f y sea para cada i $J_i = f(I_i)$. El conjunto $S - \bigcup I_i = D$ es el conjunto finito de puntos de discontinuidad. El conjunto $E = S - \bigcup J_i$ es una unión finita de intervalos, y tiene medida nula, luego es un conjunto finito de puntos. S se pone por una parte como suma conexa (ver 2.4.2.) de todos los intervalos de $\{I_i\}$ y de D , y por otra parte, como suma conexa de los intervalos de $\{J_i\}$ y de E . La diferencia entre el número de abiertos y cerrados de cada una de estas dos colecciones de intervalos debe coincidir. Como f es homeomorfismo en cada I_i , la diferencia entre el número de abiertos y cerrados en $\{I_i\}$ y en $\{J_i\}$ coincide. Por tanto debe coincidir dicha diferencia en las familias D y E — consideradas aquí como colecciones de intervalos cerrados —, y eso significa que en D y en E hay el mismo número de puntos. Es claro ahora que definiendo convenientemente las imágenes de los puntos de D sobre puntos de E , se obtiene la aplicación g deseada. ■

En una segunda etapa, probaremos que definiendo g de manera cuidadosa, puede obtenerse además la continuidad lateral en todos los puntos, aunque ahora será tal vez preciso añadir o quitar algunos extremos de las componentes conexas de S . Dividiremos el estudio en dos partes:

a) f es orientable.

Entenderemos que f es orientable si su restricción a cualquier intervalo isométrico de diámetro positivo es creciente. Sea $\{J_i\}$, $i \in M$, la familia de los intervalos isométricos de f , y definamos $H_i = \text{Int}_S(J_i)$ $\forall i \in M$. Consideremos un intervalo S' cerrado por la izquierda y abierto por la derecha, conteniendo a S estrictamente y definamos $g: S' \rightarrow S'$ de la siguiente forma: En el interior de $S' - S$, g actúa como la identidad. Si $x \in H_i$, $i \in M$, $g(x) = f(x)$. Finalmente, en el resto de los puntos, que es un conjunto finito de puntos aislados (porque son extremos de intervalos H_i no vacíos), se define g de manera que sea continua por la derecha. La imagen por g de dos puntos aislados de este último grupo, no puede coincidir, ya que a la derecha de cada uno de los dos yace un intervalo $f(H_i)$ abierto y de diámetro positivo. Esto garantiza la inyectividad de g . También debe ser sobreyectiva, ya que el complementario de la imagen de g debe ser de medida nula y se obtiene como unión finita de intervalos, es decir, debe ser un conjunto finito de puntos. Así S' se pone como unión de una familia finita y disjunta de semicerrados, que es $g(S')$, y de un número finito de puntos, que son intervalos cerrados, lo que es incompatible con ser S' semicerrado, salvo si dicha familia de puntos es vacía (ver lo referente a la suma conexa de intervalos en 2.4.2.).

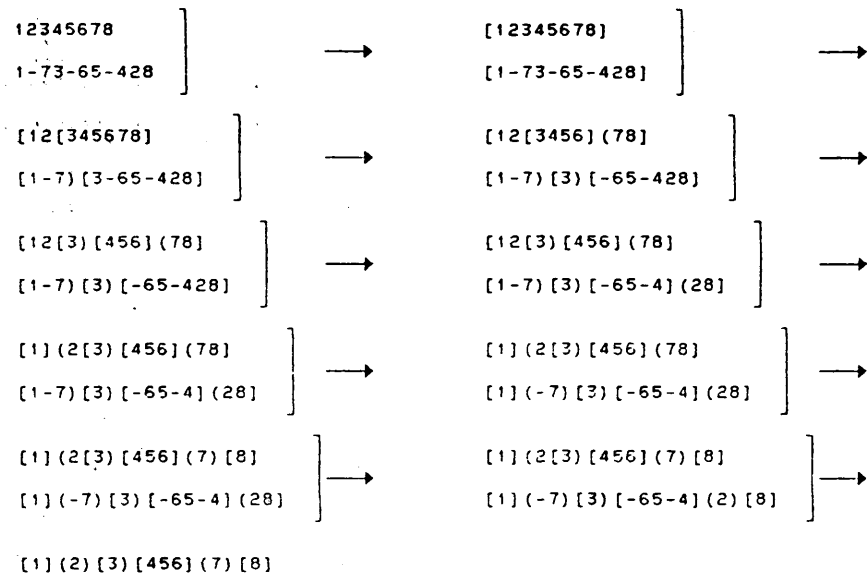
b) Caso no orientable.

Aquí la construcción es algo más complicada, aunque se basa en la misma idea. Sea S' un intervalo que incluye estrictamente a S . Supongamos que S' es un intervalo cerrado cuyo interior contiene a S , y definimos g como la identidad en el interior de $S' - S$, y g coincidente con f en los H_i definidos como en el caso anterior.

Tenemos la siguiente situación: Entre los interiores de los H_k y los interiores de los intervalos que de $S' - S$ forman una colección $\{I_k\}$ de intervalos, cuyas adherencias unidas son precisamente S' . Debemos definir g en los extremos de estos intervalos de manera que sea en ellos continua a izquierda o derecha, y que sea biyectiva. Las imágenes de los extremos inferior y superior de S' serán claramente ellos mismos. Para el resto de los extremos se procede como sigue:

Se comienza eligiendo arbitrariamente un extremo, por ejemplo el izquierdo de el intervalo I_j . Llamémoslo $x_j(i)$. A este extremo se le asigna imagen de manera que g sea en él continua por la derecha. A continuación se considera la imagen $g(I_j)$. Si g es creciente en su restricción a I_j entonces $g(x_j(i))$ será el extremo izquierdo de $g(I_j)$, pero si es decreciente, será el extremo derecho. En cualquiera de los dos casos, en el correspondiente lado de I_j yace la

imagen de un cierto l_k , $g(l_k)$, que debe ser abierta por el lado que limita con $g(l_j)$. Por tanto l_k se toma abierto por el mismo extremo que $g(l_k)$ si la restricción de g a l_k es creciente, y por el extremo contrario si es decreciente. Esto significa que el intervalo l_j que limita con l_k por el correspondiente extremo ha de ser cerrado, y así continúa el proceso hasta haber determinado la imagen de todos los extremos, o hasta retornar a $x_j(i)$. En este último caso se vuelve a elegir arbitrariamente otro extremo y a recomenzar el proceso que debe concluir en número finito de pasos puesto que el número de extremos es finito. A continuación mostramos con un ejemplo las distintas fases del proceso. En los renglones superiores están nombrados con números los intervalos. En los inferiores, aparecen sus imágenes con los mismos números, afectados con signo negativo, en el caso en que la restricción de g sea decreciente. Los intervalos 1,5,8, corresponden a $S'-S$. El proceso comienza determinando arbitrariamente que f es continua en el extremo izquierdo del intervalo 3 por la derecha (en la anterior etapa se cierran los puntos extremos de S').



Y aquí se cierra el primer ciclo comenzado al hacer g continua por la derecha en el extremo izquierdo del intervalo 3. Se volvería a recomenzar el proceso, por ejemplo, haciendo g continua por la derecha en el extremo izquierdo del intervalo 5, hasta terminar de asignar a cada extremo su imagen.

Una vez obtenida $g : S' \longrightarrow S'$, podemos retirar los intervalos auxiliares introducidos en que g actúa como la identidad (los

intervalos 1,5,8, en nuestro ejemplo). Concretamente, si retiramos las componentes isométricas de los interiores de estos intervalos, y llamamos nuevamente S' al dominio así obtenido, obtenemos el siguiente resultado, válido tanto para el caso orientable como para el no orientable:

i) Si f es una r.i. generalizada en S , hay un conjunto S' que se obtiene de S añadiendo o quitando un número finito de los extremos de las componentes conexas de S , y una reordenación por intervalos $g : S' \rightarrow S'$ que coincide con f en todos los puntos de S , salvo tal vez en el conjunto finito de los puntos de discontinuidad de f . Los puntos de cambio de g son puntos de discontinuidad de f . Además, si f no tiene puntos de discontinuidad cíclicos, $\forall x \in S$, podemos encontrar un natural k tal que $w_f(x) = w_g(f^k(x)) \cap S$.

Es en el sentido de la última aserción que la dinámica de f está subordinada a la de g . Por otra parte la hipótesis de que no haya puntos de discontinuidad de f que sean cíclicos, no es una restricción, puesto que podemos retirar el ciclo entero y estudiar la restricción de f al resto.

De todo lo afirmado sólo la última aserción está por demostrar. Pero es inmediato que si no hay puntos de discontinuidad de f que sean cíclicos, para todo x de S se puede encontrar un $k \in \mathbb{N}$ tal que si $n \geq k$ $f^n(x)$ no es de discontinuidad de f y pertenece a S , de manera que $w_f(x) = w_f(f^k(x)) = w_g(f^k(x)) \cap S$.

En el caso orientable, es fácil probar la semirregularidad de g , y podrían utilizarse los teoremas de recurrencia y estructura para estudiar la dinámica de g . Sin embargo, en el caso no orientable, no puede asegurarse que f sea semirregular, por lo que es preciso obtener generalizaciones de estos teoremas que los hagan aplicables para r.i. no semirregulares. A ello se destina el último punto de este capítulo.

2.7.2. Generalización de los teoremas de recurrencia y estructura a r.i. no semirregulares

En la demostración del teorema de recurrencia, se utilizaba la hipótesis de semirregularidad tan sólo para conseguir que el punto $z = f^p(x)$ del que se deseaba probar $z \in w(x)$ tuviera la propiedad de que para cualquier natural k fuera posible encontrar un intervalo k -isométrico de diámetro positivo conteniendo a z . Ahora bien, si f no es semirregular y suponemos que entre los puntos de cambio no hay ninguno cíclico-suprimiendo el ciclo entero, si lo hubiera-, podemos encontrar un natural K verificando que si $p \geq K$, para $z = f^p(x)$ existe tal sistema de intervalos. En efecto, para que un punto z no verifique dicha condición, es necesario que una iteración f^k sea discontinua a derecha o izquierda en z , mientras que otra iteración, f^q , sea discontinua en z en el lado contrario.

En caso en que $z \in D = S - \bigcup_{i=0}^{\infty} f^{-i}(X(f))$, z no puede estar en tal situación.

También se deduce si $z \in S \setminus D$ que si es K el mayor natural tal que $f^K(x) \in X(f)$, entonces si $p \neq K$, el punto $z = f^p(x)$ tampoco puede estar en tal situación, por lo que podemos encontrar los intervalos K -isométricos necesarios para probar que $z \in \omega(x)$. Hemos probado así el siguiente resultado, que es una generalización del teorema de recurrencia para una r.i. cualquiera:

i) Si f es una r.i. en S , para todo $x \in S$, salvo para una colección numerable de puntos, $f^p(x) \in \omega(x) \forall p \neq 0$. Para todo $x \in S$, $\exists K$ tal que si $p \neq K$, $f^p(x) \in \omega(x)$.

Como para el teorema de estructura no se precisaba la hipótesis de semirregularidad mas que para garantizar que los conjuntos $\omega(x)$ son infinitos aplicando el teorema de recurrencia, y esto resulta cierto como consecuencia inmediata del resultado anterior, para una r.i.a. cualquiera, para todo x del dominio, tenemos:

ii) Si f es una r.i.a. en S , para todo $x \in S$, $\omega(x)$ es unión finita de intervalos de diámetro positivo.

Finalmente si la r.i. no es acíclica, obtenemos:

iii) Si f es una r.i. en S , se mantiene el resultado anterior pudiendo ser ahora los intervalos que componen $\omega(x)$ de diámetro nulo. Son dos los casos posibles: Que todos los intervalos se reduzcan a puntos — caso en que $\omega(x)$ es finito —, y que todos los intervalos tengan diámetro positivo — caso en que $\omega(x)$ es infinito —.

CAPITULO III

BILLARES POLIGONALES RACIONALES

3.0 INTRODUCCION.

En este capítulo se estudia un modelo natural de aplicación de las isometrías locales biyectivas de la recta : Los billares poligonales de ángulos racionales.

En el apartado 3.1 se construye el modelo matemático para el billar poligonal. En el apartado 3.2 se estudian las propiedades del grupo R generado por las simetrías respecto a los lados del billar, y se introduce la definición de billar poligonal racional, que depende precisamente de dicho grupo. En el apartado 3.3 se expone el llamado método de la mesa móvil, consistente en suponer que la bola de billar se desplaza indefinidamente de forma rectilínea, mientras que la mesa de billar se desplaza por sucesivas simetrías respecto de sus lados. En este apartado se desarrolla el punto de vista de A. N. Zemlyakov y A. B. Katov para el estudio de los billares poligonales racionales, conectando este problema con el estudio de flujos geodésicos en superficies orientables. En el apartado 3.4 se estudian propiedades del subgrupo de traslaciones T del grupo R , mostrándose su conexión con las órbitas periódicas en el apartado 3.5. En el apartado 3.6 se prueban los resultados sobre minimalidad en caso de órbitas de orientaciones genéricas, estructura de una órbita cualquiera y caracterización de orientaciones singulares y finalmente en el anexo de gráficos de este capítulo se muestran diseños obtenidos a partir de los conjuntos A_p descritos en 3.4.3, y órbitas periódicas obtenidas por el procedimiento descrito en 3.5.3

3.1. MODELO DE BILLAR POLIGONAL.

3.1.1. La mesa del billar.

Consideremos una curva de Jordán poligonal plana limitando una región Q del plano. No se requiere que la poligonal sea convexa, y además consideramos la posibilidad de la existencia de un conjunto finito de segmentos cerrados incluidos en Q - los obstáculos -. llamaremos a Q mesa del billar. Al conjunto de segmentos de la poligonal y de segmentos incluidos en Q lo llamaremos borde del billar y lo denotaremos $\partial Q = \{ l_i \}$, $1 \leq i \leq n$. Los l_i son los lados del billar, los extremos de los lados forman el conjunto V de los vértices del billar. Un vértice puede pertenecer a dos lados. El vértice común a los lados l_i y l_j lo denotaremos v_{ij} . Escribiremos α_{ij} para indicar el ángulo formado por los lados l_i y l_j . A denota el conjunto de todos los ángulos.

3.1.2. Modelo físico para el billar.

Estudiamos las trayectorias de puntos materiales que se mueven con movimiento rectilíneo en Q , rebotando en ∂Q según la ley de la reflexión elástica, es decir, con igual ángulo de incidencia y reflexión, y de forma indefinida.

Con el fin de describir matemáticamente el citado sistema dinámico, que se presenta como modelo físico, construiremos un espacio de fases adecuado cuyos puntos representarán los puntos materiales en movimiento rectilíneo, y una bivección en el espacio de fases, cuyas iteraciones sobre cada punto del espacio de fases sean los sucesivos tramos de la trayectoria que describiría la partícula material a la que identificamos el punto del espacio de fases.

3.1.3. Espacio de fases y correspondencia de reflexión.

Consideramos las semirrectas con origen en ∂Q . Una de tales semirrectas puede ser descrita como un par ordenado (P, θ) cuyo primer elemento es un punto $P \in \partial Q$ y cuyo segundo elemento es un ángulo $\theta \in R/2\pi Z$ que expresa el ángulo que forma la semirrecta con cierta semirrecta OX con origen en un punto arbitrario O y que actúa de semirrecta origen para la medida de ángulos. Nuestro espacio de fases será el conjunto Q de aquellas semirrectas (P, θ) con $P \in \partial Q$ y tales que P es adherente a $\text{Int}(Q) \cap (P, \theta)$. Si el recinto Q fuera convexo, esta última condición se reduciría a la exigencia de que (P, θ) cortara al interior de Q . Definimos en Q una correspondencia de reflexión τ haciendo corresponder a cada $(P, \theta) \in Q$ una $\tau(P, \theta) \in Q$ de la siguiente manera: Si $P' \neq P$ es el punto de $(P, \theta) \cap \partial Q$ más próximo a P y P' pertenece al lado l_i de Q , llamando s_i a la simetría axial respecto de l_i , definimos $\tau(P, \theta) = s_i(P', \theta)$. Nótese que para que esta correspondencia esté bien definida, es preciso que P' pertenezca a un único l_i . Si P' fuera un vértice, no quedaría bien determinado el lado respecto del que debemos tomar la simetría. Por el momento consideraremos estos casos excepcionales. Más adelante nos ocuparemos de resolver satisfactoriamente el problema de los rebotes en los vértices. Provisionalmente, y para que τ sea una aplicación, podemos poner $\tau(P, \theta) = (P, \theta)$ si $P' \in V$. Hemos definido así una aplicación $\tau: Q \rightarrow Q$ tal que dado un elemento $(P, \theta) \in Q$, la sucesión $\{\tau^n(P, \theta)\}$, $n \in N$ — a cuyos términos llamaremos elementos de trayectoria —, describe de forma adecuada la trayectoria de nuestra partícula material rebotando elásticamente contra los lados del billar. (Con más precisión la trayectoria está formada por una sucesión de segmentos $\tau^n(P, \theta) \cap Q$, pero no se utilizarán en general estos elementos de trayectoria "reales" por complicar la notación).

Sea dx la medida de Lebesgue en ∂Q , y sea $d\theta$ la medida de Haar en $R/2\pi Z$. Hay una medida natural en Q como subespacio de $\partial Q \times R/2\pi Z$, dada por $dx \times d\theta$, aunque en 3.6 veremos que resulta más interesante una modificación de esta definición para lograr que τ conserve la medida. De momento consideraremos esta medida en Q .

Llamaremos billar poligonal al sistema dinámico formado por Q y τ , y lo denotaremos (Q, τ) . Las definiciones para O^+ , O^- , O , y para los operadores α y ω son las obvias (ver 2.1.1. y 2.6.1).

3.2. GRUPO GENERADO POR ∂Q

3.2.1. Definición.

Dada una semirrecta arbitraria $(P, \theta) \in Q$, escribiremos $[(P, \theta)]$ para representar la recta que contiene a (P, θ) . Llamemos Γ al conjunto de todas las rectas $[(P, \theta)]$ cuando (P, θ) recorre Q . La misma correspondencia de reflexión definida en Q puede extenderse al conjunto Γ si hacemos $\tau[(P, \theta)] = [\tau(P, \theta)]$. Además, $\tau(P, \theta) = s_i(P', \theta)$ donde P' es cierto punto de ∂Q (ver 4.1.3.) y s_i es una simetría axial respecto a cierto lado de ∂Q , y es inmediato comprobar que en estas condiciones es $\tau[(P, \theta)] = s_i[(P, \theta)]$. Para $\tau^n[(P, \theta)]$ se obtiene:

$$\tau^n[(P, \theta)] = g[(P, \theta)]$$

Donde g es un producto de simetrías axiales con ejes en los lados de ∂Q . Puesto que la trayectoria $\{\tau^n(P, \theta) \cap Q\}$ está contenida en el conjunto $\{\tau^n[(P, \theta)]\}$, resulta natural estudiar la operación del grupo generado por las simetrías axiales de ejes en los lados de ∂Q sobre el conjunto Γ , y aprovechar las propiedades de dicho grupo.

Así pues definimos el grupo generado por ∂Q como el grupo R generado por el conjunto $\{s_i\}$, $1 \leq i \leq n$, donde s_i es la simetría axial respecto al lado l_i de ∂Q .

3.2.2. Estructura de R

El grupo R admite como subgrupo normal el subgrupo T de traslaciones de R . Si llamamos H al grupo cociente R/T y $g_1, g_2 \in R$ son dos transformaciones equivalentes para la relación de equivalencia inducida en R por T , si consideramos una semirrecta cualquiera (P, θ) , las semirrectas $g_1((P, \theta))$ y $g_2((P, \theta))$ tienen la misma orientación, por lo que el número de distintas orientaciones que pueden adoptar las imágenes por elementos de R de una semirrecta dada, no puede exceder al número de clases de equivalencia en R , es decir, al orden del grupo H . De aquí el interés en obtener una representación isomorfa de H que nos permita conocer la estructura de este grupo.

Para ello establecemos el homomorfismo de grupos siguiente: A cada lado $l_i \in \partial Q$, hacemos corresponder una simetría axial s'_i de eje paralelo al s_i y que pasa por un punto determinado O independiente de i , de manera que la simetría s_i de eje l_i se relaciona con s'_i por la relación: $s_i = t \circ s'_i \circ t^{-1}$, siendo t cierta traslación (la que transporta l_i sobre el eje de s'_i). Las s'_i , como son simetrías axiales que pasan todas por el mismo punto O , generan un grupo diedral D , subgrupo del grupo de los movimientos que dejan O invariante. Nos proponemos probar:

$$i) H \cong D$$

La correspondencia $\tilde{\phi} : R \longrightarrow D$ definida por:

$$\tilde{\phi}(s_1 \cdot s_2 \cdot s_3 \cdot \dots \cdot s_k) = s'_1 \cdot s'_2 \cdot s'_3 \cdot \dots \cdot s'_k$$

Define un epimorfismo de grupos. Para ver que se trata de una aplicación, probaremos primero que cualquier expresión de la identidad como producto de elementos de R , tiene como imagen por $\tilde{\phi}$ la identidad. Para ello observamos que, para cualquier semirrecta (P, θ) , la orientación de $s_i((P, \theta))$ coincide con la de $s'_i((P, \theta))$. Como consecuencia de ello un producto de simetrías de R y su imagen por $\tilde{\phi}$ en D transforman cualquier semirrecta en dos semirrectas de la misma orientación. Si tal producto es la identidad, su imagen por $\tilde{\phi}$ en D transforma cualquier semirrecta en otra de su misma orientación. Pero las transformaciones en el grupo diedral D son giros o productos de un giro por una simetría de D , y la única transformación de este tipo que conserva la orientación de cualquier semirrecta es la identidad. De aquí se deduce que la imagen por $\tilde{\phi}$ de la identidad es la identidad. Con este resultado, se obtiene fácilmente que dos expresiones distintas como productos de elementos de R de una misma transformación, tienen imagen igual por $\tilde{\phi}$, de donde $\tilde{\phi}$ es una aplicación. La conservación de la estructura de grupo por $\tilde{\phi}$ es inmediata, y también que se trata de una aplicación sobreyectiva.

El mismo razonamiento aplicado más arriba para comprobar que $\tilde{\phi}(\text{id}) = \text{id}$, sirve para comprobar que si $g \in R$ conserva la orientación de cualquier semirrecta, $\tilde{\phi}(g) = \text{id}$, y recíprocamente, si g altera la orientación de una semirrecta, su imagen por $\tilde{\phi}$ no será la identidad. De ello se deduce que el núcleo de $\tilde{\phi}$ son las transformaciones de R que no alteran la orientación de ninguna semirrecta, es decir el subgrupo T de traslaciones de R . Ello demuestra el isomorfismo de grupos $H = R/T \cong D$. ■

3.2.3. Billares poligonales racionales

Cuando el grupo $H \cong D$ es finito, decimos que el billar (Q, τ) es un billar poligonal racional, o más brevemente, un billar racional. La principal propiedad de los billares racionales es:

i) Si (Q, τ) es un billar racional, el conjunto de las orientaciones de los elementos de trayectoria de un (P, θ) arbitrario de Q , es un conjunto finito.

En efecto, sabemos que $\tau^n((P, \theta))$ es cierta semirrecta contenida en la recta $q_n[(P, \theta)]$, con $q_n \in R$. Su orientación es una de las dos posibles orientaciones de esta, y de hecho se prueba fácilmente que coincide con la de $q_n((P, \theta))$, y como se razonó en el anterior punto las orientaciones de las semirrectas de este tipo, no pueden exceder el cardinal de H . Este hecho nos permitirá más adelante simplificar de forma decisiva el espacio de fases del problema del billar. Por ahora nos ocuparemos de las condiciones bajo las cuales puede afirmarse que un billar es racional.

El grupo diedral D generado por simetrías respecto ejes s'_i que se cortan en un punto O , se pone como producto indirecto:

$$D = G \times S$$

Donde G es el subgrupo normal de los giros de D , y S es el grupo de orden dos generado por una simetría s'_i cualquiera. Se trata de un producto indirecto, puesto que S no es un subgrupo normal.

A su vez el subgrupo G , coincide con el grupo generado por los productos de un número par de las s'_i . Está generado por tanto por transformaciones de la forma $s_i \circ s_j$, que a su vez son giros de centro en O y ángulos $2\alpha_{ij}$. Es claro que si alguno de estos ángulos es irracional (se expresa como una fracción irracional de 2π), el subgrupo de G generado por $g(O, 2\alpha_{ij})$ no sería finito (sería de hecho denso en el grupo de los giros de centro en O , según el teorema de Kronecker). Recíprocamente, si todos los α_{ij} son racionales, es fácil ver que G coincide con un grupo cíclico de orden p , donde p es el mínimo común múltiplo de los denominadores de las fracciones que expresan los $2\alpha_{ij}$ como fracciones irreducibles de 2π , o más simplemente, que expresan los α_{ij} como fracciones irreducibles de π . A su vez, $D = G \times S$ resulta ser en este caso el grupo diedral D_p asociado a un polígono regular de p lados, centrado en O , y tal que alguno de sus ejes de simetría coincide con alguno de los s'_i . Este grupo tiene orden igual a $2p$. Resumiendo obtenemos:

ii) La condición necesaria y suficiente para que (Q, r) sea racional, es que todos los ángulos α_{ij} sean fracciones racionales de π . Si el mínimo común múltiplo de los denominadores de dichas fracciones, supuestas irreducibles, es p , $H \cong D_p$.

En el caso particular en que ∂Q sea conexo se obtiene:

iii) Para que (Q, r) sea racional, es suficiente que los ángulos de cada vértice de Q , sean racionales.

En efecto, es fácil probar por inducción que si esta condición se cumple, entonces también son racionales los ángulos formados por cualquier pareja de lados no consecutivos. Sin embargo, en el caso de billar con obstáculos, esta condición no es suficiente.

3.3 EL METODO DE LA MESA MOVIL.

En este apartado se describe un interesante punto de vista alternativo al desarrollado en el anterior apartado, consistente en considerar el movimiento de la bola de billar indefinidamente rectilíneo, y paralelamente una mesa móvil que se desplaza junto con la bola. Este punto de vista aporta muchos resultados interesantes más fáciles de ver que por el método de las reflexiones de la bola.

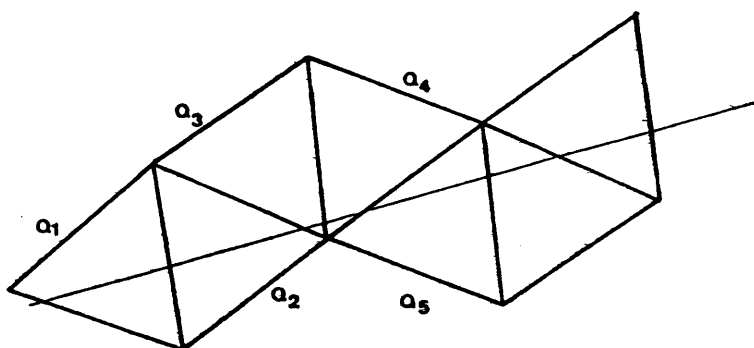


fig. 3.1

3.3.1.

Sea (Q, τ) un billar poligonal en una mesa Q de borde ∂Q . Tomamos la semirrecta $(P, \theta) \in Q$, y supongamos que la primera reflexión de la misma tiene lugar en el lado l_i . Consideremos el elemento de trayectoria $\tau(P, \theta) = s_i(P_1, \theta)$, donde s_i es la simetría axial respecto a l_i y $P_1 = (P, \theta) \cap l_i$. Consideremos también la región $Q_1 = s_i(Q)$ y en ella un billar (Q_1, τ) de borde $\partial Q_1 = s_i(\partial Q)$. El elemento de trayectoria $\tau(P, \theta)$ se transforma por sí en un elemento de trayectoria de (Q_1, τ) igual a $s_i \tau(P, \theta) = s_i s_i(P_1, \theta) = (P_1, \theta)$, de manera que (P_1, θ) está en la misma recta que (P, θ) . Aplicando el mismo razonamiento a (P_1, θ) en (Q_1, τ) obtenemos una sucesión de elementos de trayectoria $\{(P_n, \theta)\}$ pertenecientes a una misma recta, donde (P_n, θ) es un elemento de trayectoria en un billar (Q_n, τ) de mesa Q_n y borde ∂Q_n obtenidos para cada n por reflexión en el correspondiente lado del anterior billar. Es decir se procede como si la bola prosiguiera su trayectoria rectilínea, mientras que la mesa se transforma por reflexiones sucesivas (ver fig. 3.1).

3.3.2.

Nos preguntamos ahora cual es la relación que existe entre las mesas Q y Q_n tras n reflexiones. El resultado que vamos a probar es el siguiente:

i) Si tras n reflexiones es $Q_n = g_n(Q)$, y en el billar (Q, τ) , (P, θ) se transforma en $\tau^n(P, \theta)$, se tiene:

$$g_n = (\tau^n)^{-1}$$

Si el rebote n -simo de (P, θ) en Q es sobre el lado s_j , el elemento de trayectoria (P_n, θ) del billar (Q_n, τ) incidirá sobre el lado $g_n(l_j)$. La simetría con este eje se escribe $g_n \circ s_j \circ (g_n)^{-1}$, de donde $g_{n+1} = g_n \circ s_j \circ (g_n)^{-1} \circ g_n = g_n \circ s_j$. Por otra parte, en el billar (Q, τ) tenemos que $\tau^{n+1} = s_j \circ \tau^n$. Esto prueba que g_{n+1} y τ^{n+1} se obtienen componiendo las mismas simetrías en orden inverso, y por tanto que para todo n se verifica $g_n = (\tau^n)^{-1}$, que era la relación que buscábamos. ■

3.3.3 Billares racionales y flujos geodésicos en superficies.

Supongamos ahora que el billar (Q, τ) es racional, y que el borde ∂Q es conexo. Dada un elemento inicial de trayectoria (P, θ) , establecemos en la sucesión $\{Q_n\}$ de mesas una relación de equivalencia, cuyas clases están compuestas de mesas que pueden obtenerse mutuamente por traslaciones. Como el grupo H es finito, es claro que el número de clases de equivalencia es finito e igual al orden m de H . Formamos un conjunto C compuesto por una mesa de cada clase de equivalencia. Sea entonces $C = \{M_i\}, 1 \leq i \leq m$, donde cada M_i es cierta mesa de la sucesión $\{Q_n\}$. Llamemos ahora l_{ij} al j -ésimo lado de la mesa M_i . Si s_{ij} expresa la simetría axial respecto a dicho lado, sabemos que cierta mesa M_k de C es tal que $t s_{ij}(M_i) = M_k$, donde t es una traslación, y tal M_k es única en C . Procedemos ahora a pegar el lado l_{ij} de la mesa i al lado $t s_{ij}(l_{ij}) = t(l_{ij}) \in M_k$, identificando cada punto x de l_{ij} al punto $t(x)$. Si esta operación se realiza con todos los lados de todos los polígonos de C , se obtiene una superficie M compacta y orientable. Al elemento de trayectoria $\tau^n(P, \theta)$ en el billar (Q, τ) corresponde el elemento (P_n, θ) en el billar (Q_n, τ) . Mediante la proyección canónica de las $\{Q_n\}$ sobre C , a (P_n, θ) corresponde un elemento de trayectoria en cierta mesa M_i de C , y a este corresponde también canónicamente un segmento en M . Como todos los (P_n, θ) tienen la misma orientación θ , la unión de sus proyecciones canónicas sobre M será una geodésica sobre la superficie M . Inversamente, a una geodésica en M corresponde la trayectoria de una bola en el billar (Q, τ) . Esta bonita construcción, que pone de manifiesto la relación existente entre billares racionales y flujos geodésicos en superficies es debida a A.N.Zemlyakov y A.B.Katov.

Como ya hemos señalado, la superficie M es orientable, por lo que será en general una esfera con k asas. Si Q es un rectángulo, entonces M es un toro. En la figura 3.2 se muestra el caso en que Q

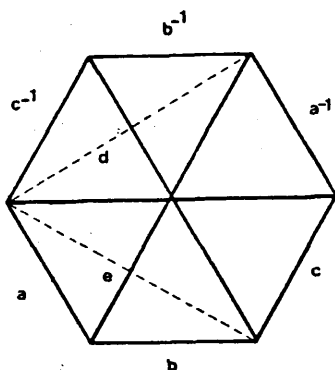


fig. 3.2

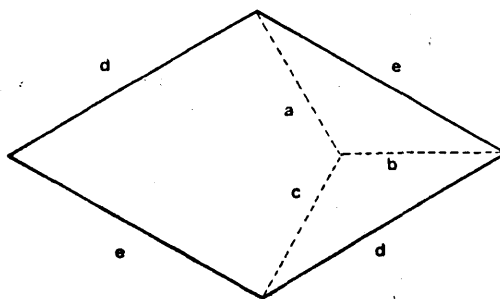


fig 3.3

es un triángulo equilátero. En este caso, el grupo D_3 es el grupo diedral asociado a un triángulo equilátero, de orden seis. Las seis mesas correspondientes se pueden pegar de forma que completan un exágono, como se indica en la figura. El resto de las identificaciones necesarias para obtener la superficie M están indicadas poniendo a los lados que deben identificarse nombres iguales. La superficie resultante se puede expresar mediante la sucesión de los lados que completan el borde del exágono recorrido en este caso en sentido positivo. Las flechas de los lados expresan como se deben identificar al lado de su mismo nombre (origen con origen y extremo con extremo). Si un lado resulta recorrido en sentido contrario al suyo propio, se indica en la sucesión mediante el exponente -1 . Así, la superficie M se representa por $abca^{-1}b^{-1}c^{-1}$. Según las técnicas clásicas utilizadas para la demostración del teorema de clasificación de superficies, los pares de aristas de primera especie aa^{-1} y bb^{-1} que se separan mutuamente, pueden ser escritos de forma consecutiva, resultando $M = aba^{-1}b^{-1}cc^{-1}$, y el par de aristas de primera especie consecutivas cc^{-1} se puede eliminar con lo que M se escribe: $aba^{-1}b^{-1}$, que es la forma normal de un toro. A este resultado se puede llegar también por el método indicado en figuras 3.2 y 3.3, en donde se indica respectivamente como cortar y como pegar el exágono para obtener un toro.

En la figura 3.4 se estudia el caso en que Q es un triángulo rectángulo de ángulo $\pi/8$. En este caso D_8 tiene orden 16. La colección de 16 triángulos de T se puede pegar formando el octógono de la figura. M se escribe: $abcd a^{-1}b^{-1}c^{-1}d^{-1}$. Los pares de primera especie aa^{-1} y bb^{-1} , mediante las oportunas transformaciones, pueden ser escritos de forma consecutiva, y M se escribe $aba^{-1}b^{-1}cdc^{-1}d^{-1}$, que es la forma normal de la suma conexa de dos toros, o bien de una esfera con dos asas. En la misma figura se indica la construcción de los primeros elementos de trayectoria para un (P, Θ) inicial.

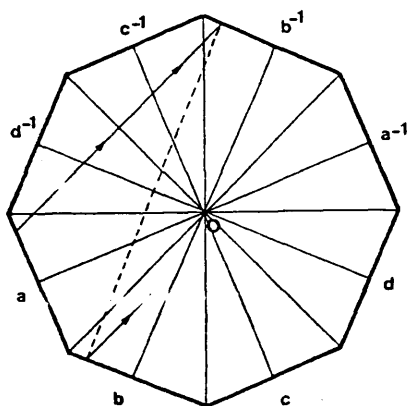


fig 3.4

Este último ejemplo muestra como un triángulo rectángulo de ángulo $\pi/4k$, da lugar a una superficie M que es una esfera con k asas.

En el caso particular en que la superficie M resulte ser un toro, podemos reducir el problema del billar en la mesa Q , al estudio del flujo geodésico en el toro, problema cuya solución es bien conocida. Particularmente, el teorema de Kronecker bidimensional, asegura que si tal flujo no es periódico - se suele llamar irracionales a este tipo de flujos toroidales -, entonces es minimal, y por consiguiente también lo será el billar en el correspondiente Q . Un teorema de Weyl (ver refs [H1] y [W1]) permite asegurar que los flujos irracionales en un toro y en consecuencia los billares correspondientes, son ergódicos. Los ya mencionados casos en que Q es un rectángulo o un triángulo equilátero, junto con los triángulos rectángulos de ángulos $\pi/4$ y $\pi/6$, y el exágono regular (obsérvese : todos los polígonos que teselan el plano por simetrías), están en este caso.

Sin embargo, el ya citado resultado de Gal'perin (ver introducción a la tesis y ref. [G.1]) prueba que este resultado no puede ser generalizado a cualquier billar racional.

3.4 EL SUBGRUPO T DE TRASLACIONES

En un billar poligonal cualquiera, el grupo $H = R/T \cong D$, es un grupo finitamente generado, y sin embargo T no lo es en general. En el caso en que el billar sea racional, en que sabemos que D_p es finito, T no lo es, aunque como vamos a comprobar, en este caso si es finitamente generado. En este apartado supondremos que ∂Q es conexo.

3.4.1 Estructura del subgrupo de traslaciones T

Sea $L = \{ l_i, i = 1, \dots, n \}$ la colección de todos los lados de un billar, considerados ahora como vectores planos. Es fácil ver que el conjunto $RL = \{ g(l) : g \in R, l \in L \}$ genera un grupo $\langle RL \rangle$ que incluye a T . En efecto, toda t de T , puede ser obtenida como producto de simetrías s_i respecto a los lados de L . Un razonamiento similar al desarrollado en 3.3.2 muestra que si h_k es la transformación que se obtiene componiendo las k primeras simetrías que intervienen en t en orden inverso, entonces las $Q_k = h_k(Q)$ forman una sucesión de mesas en la que Q_k se obtiene como simétrica de Q_{k-1} respecto de un lado de esta última (la diferencia con 3.3.2 es que entonces se elegía siempre el lado de Q_{k-1} en el que incidía la prolongación de cierto elemento de trayectoria inicial, y ahora se pueden elegir otros lados). Si el número total de simetrías que componen t es m , h_m será precisamente t^{-1} . Una cadena de lados de los distintos Q_k conducen desde un vértice cualquiera $A \in Q$ al

vértice $t^{-1}(A) \in Q_m$. Cada uno de los lados de cada Q_k pertenece al conjunto RL , ya que cada Q_k es igual a $h_k(Q)$, con $h_k \in R$. Y la suma de todos los vectores de la cadena, es igual a t^{-1} . Esto prueba que $\langle RL \rangle$ incluye a T .

Podemos afinar más, eliminando de L dos lados consecutivos cualesquiera se obtiene una colección L' y la $RL' = \{g(l) : g \in R, l \in L'\}$ también verifica $T \subset \langle RL' \rangle$, y dos es además el máximo número de lados de L que se pueden suprimir para obtener el resultado. En efecto, dos es el número máximo de lados consecutivos del borde de un polígono K que se pueden suprimir de manera que si s es la simetría respecto de un lado cualquiera de K , el borde de $K \cup s(K)$ siga siendo conexo (observar en la figura 3.5 que también conexo cuando la simetría es sobre los lados suprimidos. Los lados marcados en grueso en dicha figura, corresponden a los lados de L' y a sus transformados por R .)

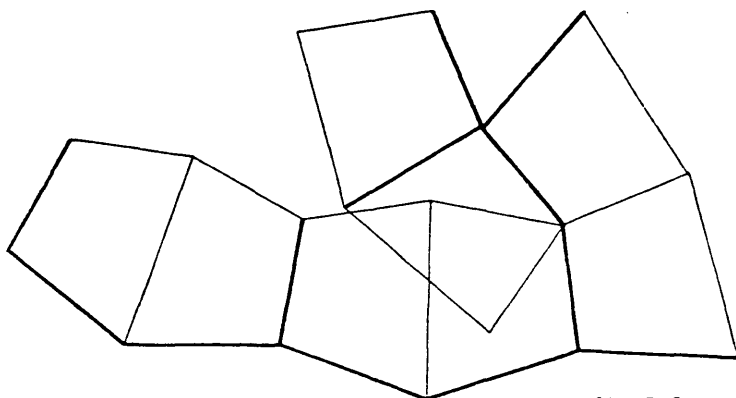


fig 3.5

Observemos ahora que si $g \in R$ y $\tilde{g}(q) = h \in D$ (ver 3.2.2 para la definición de \tilde{g} y D), para cualquier $l \in L'$ será $g(l) = h(l)$, ya que sabemos que g y h son movimientos del plano que transforman cualquier semirrecta en dos semirrectas de la misma orientación, y por tanto sus imágenes respectivas de l serán iguales como vectores libres del plano. El grupo $\langle RL' \rangle$ será por tanto igual al grupo $\langle DL' \rangle$, es decir, el grupo generado por el conjunto de vectores $DL' = \{h(l), h \in D, l \in L'\}$. Se obtiene:

i) Si el billar es racional, el subgrupo T de traslaciones de R es finitamente generado. Si el billar no es racional, T es en todo caso numerable.

Ya que en tal caso $D = D_p$ es un grupo finito, y por tanto también lo es el conjunto $D_p L'$ que genera al grupo abeliano $\langle RL' \rangle$. Como T es un subgrupo de un grupo abeliano finitamente generado, también lo será T . Si el billar no es racional, como $\langle RL' \rangle$ es en todo caso generado por un conjunto numerable, será numerable, y también T . ■

3.4.2. Casos particulares

En tres casos particulares la estructura de $\langle RL' \rangle$ es especialmente sencilla:

i) El grupo $\langle RL' \rangle = \langle DL' \rangle$, del cual T es subgrupo, es el Z -módulo engendrado por las raíces p -ésimas de la unidad en los siguientes casos: a) Q es un triángulo racional; b) Q es un polígono regular de k lados; c) Q es un polígono racional construido con lados de igual longitud.

A continuación se analizan estos tres casos, detallándose en los dos primeros el valor de p .

a) Q es un triángulo racional.

Si la mesa Q de billar es triangular, entonces L' consta de un solo vector l , que es un lado arbitrario del triángulo, por lo que DL' es el conjunto de transformados de dicho vector por el grupo D . Si el billar no es racional DL' será un conjunto numerable de vectores, denso en el conjunto de todos los vectores planos de módulo $|l|$ (teorema de Kronecker). Por el contrario, si el billar es racional, DL' es la órbita de el vector l por el grupo D_p de simetrías de un polígono regular de p lados, uno de cuyos ejes de simetría es paralelo a l (ver 3.2.3). Como l es uno de los ejes de simetría que generan D_p , dicha órbita no será otra que todos los vectores con origen en el centro y extremos en los vértices de un polígono regular de p lados (donde p era, según se dijo en 3.2.3, el mínimo común múltiplo de los denominadores que expresan los tres ángulos del triángulo como fracciones irreducibles de π). Finalmente, el subgrupo $\langle DL' \rangle$, es el grupo abeliano generado por dichos vectores. Tendremos pues que $\langle DL' \rangle$ es isomorfo al grupo abeliano A_p engendrado por las raíces p -ésimas de la unidad. De hecho, si suponemos que $|l| = 1$ y que l es horizontal y está situado en el origen, entonces $\langle DL' \rangle$ coincide con A_p .

b) Q es un polígono regular de k lados

En este caso, tendremos en cuenta que todos los lados de Q se pueden obtener de uno dado l , por simetrías respecto a las diagonales del polígono. Los elementos de DL' serán todas las todos los vectores de la forma $g \cdot h(l)$ donde g pertenece al grupo diedral D generado por simetrías respecto a ejes paralelos a los lados de Q y que pasan por el punto O , que supondremos el centro de Q , y h al grupo diedral D' generado por simetrías de ejes paralelos a las diagonales de Q y pasando por O . Por tanto $g \cdot h$ pertenecerá al grupo diedral D_p generado por las simetrías de D y D' , DL' coincidirá con D_{p1} , y $\langle D_{p1} \rangle$ es isomorfo a A_p . Se puede caracterizar exactamente el valor de p dependiendo de k .

Teniendo en cuenta que el ángulo formado por un lado y la diagonal de un polígono regular que concurre en uno de los vértices del lado vale, en fracción de 2π , $(k-2)/4k$, los ángulos formados por los ejes de simetría de D con los de D' , serán de la forma $(k-2)/4k + r/k \bmod 1$, $0 \leq r < k$, por tanto todos ellos múltiplos del ángulo

$1/4k$, por lo que $D_p \subset D_{2k}$, siendo este último el grupo diedral asociado a un polígono regular de $2k$ lados, uno de cuyos vértices coincidirá con los de Q . (tengase en cuenta que los ejes de simetría de un polígono de $2k$ lados, dividen la circunferencia circunscrita en $4k$ partes). Podemos distinguir tres casos:

$b_1)$ k es impar

Teniendo en cuenta que $k = 4s + 1$ o $4s + 3$, entre los ángulos formados por lados y diagonales, está el $1/4k$. Por tanto, el grupo generado por las simetrías de D y las de D' es el grupo diedral D_p con $p = 2k$.

$b_2)$ $k = 4s + 2$ (exágono, decágono)

Ahora los ángulos son de la forma π/k , por lo que el grupo es D_p con $p = k/2$

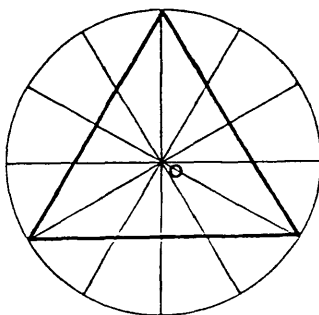


fig 3.6

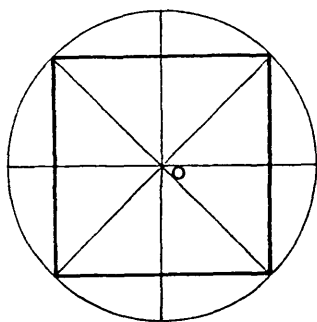
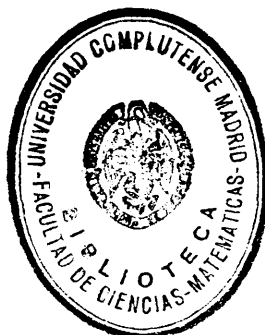


fig 3.7



b₃) $k = 4s$

Los ángulos son múltiplos de $1/2k$, por lo que el grupo es D_p con $p = k$

En las figuras 3.2 se puede ver el caso b_2 , en las figuras 3.6 y 3.7, los casos b_1 y b_3 respectivamente.

c) Q es un polígono racional de lados de igual longitud.

Un razonamiento similar al hecho anteriormente para los polígonos regulares muestra que también en este caso $\langle RL' \rangle = \langle DL' \rangle$ es isomorfo al \mathbb{Z} -módulo generado por las raíces p -ésimas de la unidad, dependiendo p en este caso del polígono Q . Un ejemplo característico de este tipo, es de los polígonos Q que se pueden construir con tramos horizontales y verticales de la misma longitud, en que $\langle RL' \rangle$ será isomorfo al conjunto $\mathbb{Z} \times \mathbb{Z}$, o bien al \mathbb{Z} -módulo A_4 generado por las raíces cuartas de la unidad.

3.4.3. Algunas propiedades de los A_p

En el anexo a este capítulo se muestran algunos de los grupos A_p , para cuya construcción se dibujan los vértices de un p -ágono regular de lado 1, y se dibujan después, tomando como centros estos vértices, nuevos p -ágonos regulares con la misma orientación y lado que los anteriores, procediéndose de igual manera a partir de los vértices de estos últimos. Se ponen de manifiesto bellas simetrías de estos conjuntos si en lugar de unir cada centro con los vértices de su p -ágono, se unen los vértices consecutivos entre sí, o bien alternativamente para formar los polígonos estrellados.

La propiedad que queremos destacar es la siguiente:

i) A_p es denso en el plano, excepto cuando $p = 1, 2, 3, 4, 6$

En el caso $p = 1, 2$, A_p es el conjunto \mathbb{Z} de los enteros. Si $p = 3, 6$, las raíces terceras o sextas de la unidad generan una red exagonal que tampoco es densa, y A_4 es una red de vértices de cuadrados, los llamados enteros de Gauss en el plano complejo.

Para probar que en el resto de los casos A_p es denso, utilizaremos el hecho de que el conjunto A_p dotado de la operación de producto de números complejos es un anillo. Si existe un número $c \in A_p$ con $0 < |c| < 1$, entonces siendo ε positivo y arbitrariamente pequeño, hay un elemento $x \in A_p$ tal que $|x| < \varepsilon$. Basta tomar $x = c^n$ para n suficientemente grande. Consideremos también $y = xe^{i\theta}$ con $\theta = 2\pi/p$. Se tiene que $y \in A_p$ y que $|y| = |x|$. Las combinaciones lineales de x e y con coeficientes enteros forman una red de rombos de diámetro menor que 2ε que cubren el plano. Esto prueba

que A_p es denso si para algún $c \in A_p$ es $1 < c < 1$. Se comprueba que $c = e^{i\theta} - 1$, con $\theta = 2\pi/p$ verifica esta condición si $p > 6$. El único caso restante, $p = 5$, también verifica la condición tomando $c = 1 - e^{i\theta} = e^{i4\theta}$.

3.5 ORBITAS PERIÓDICAS.

En este apartado se establece una condición necesaria para que una órbita en un billar poligonal sea periódica, y también se formula una conjetura apoyada por pruebas experimentales en computador, de los casos en que esta condición necesaria pudiera ser suficiente.

3.5.1.

Decimos que la órbita descrita por un elemento de trayectoria (P, θ) en un billar (Q, r) es periódica, si para algún $n \in \mathbb{N}$ es $\tau^n(P, \theta) = (P, \theta)$. Si la mesa del billar es un triángulo acutángulo, se sabe que siempre existe una órbita periódica que consta de tres elementos de trayectoria que unen los pies de las tres alturas del triángulo. También se sabe — más adelante se desprenderá de los resultados que vamos a presentar — que el conjunto de orientaciones para las cuales hay órbitas periódicas, es numerable. Pero se saben muy pocas cosas aparte de las ya mencionadas y las que presentamos a continuación. Remitimos al la ref. [K2] donde se mencionan varias cuestiones abiertas en relación a las órbitas periódicas, entre ellas, si existen órbitas periódicas en los billares triangulares obtusángulos.

3.5.2. Una condición necesaria para la periodicidad.

i) Para que la órbita descrita por el elemento de trayectoria (P, θ) sea periódica, es condición necesaria que su orientación θ sea la de alguna traslación de T .

Supongamos que (P, θ) describe una órbita periódica, y por tanto que $(P, \theta) = \tau^n(P, \theta)$. Según lo expuesto en 3.2.1 sabemos que $\tau^n[(P, \theta)] = g[(P, \theta)]$, siendo g cierta transformación del grupo R . Pero según las definiciones $\tau^n[(P, \theta)] = [\tau^n(P, \theta)] = [(P, \theta)]$, esta última igualdad por la periodicidad. Se obtiene que $g[(P, \theta)] = [(P, \theta)]$. Así, g es una transformación de R que conserva una recta invariante. De esta forma g^2 sería entonces un movimiento directo que deja fija la recta $r = [(P, \theta)]$. Por tanto puede ser únicamente una traslación o un giro de π radianes y centro en r . En tal caso, $h = g^4$ debe ser una traslación, eventualmente la identidad. Sólo resta probar que h no puede ser la identidad, ya que si no lo es h es una traslación que deja r invariante, y tiene por tanto orientación θ o $\theta + \pi$, pero en este caso, $h^{-1} \in T$ tendría orientación θ . Pero h no puede ser la identidad, ya que si suponemos que la trayectoria se prolonga siempre en línea recta, mientras que las mesas se van transformando por simetrías sucesivas sobre sus lados, según se describe en el apartado 3.3., la transformación h^{-1} es la que

transforma la mesa inicial en la mesa Q_n , y esta es una traslación de módulo igual a la longitud total de la trayectoria al cabo de n reflexiones. ■

3.5.3. Una conjetura

En el anexo a este capítulo se presentan algunos ejemplos de órbitas periódicas obtenidas eligiendo una orientación θ correspondiente a una traslación de T y estudiando la trayectoria descrita por un elemento (P, θ) . Como se ve en los ejemplos, algunas de las órbitas periódicas así halladas constan de varios cientos de rebotes, y algunas incluso de miles de ellos. También se adjunta el programa con el que han sido halladas las órbitas. Este programa permite la construcción de un billar poligonal arbitrario y el estudio de la trayectoria descrita por un elemento de trayectoria arbitrario. El programa busca primero una traslación de R , estudiando la trayectoria de un elemento (P, θ) hasta encontrar un elemento de trayectoria que tenga la misma orientación θ tras un número par de rebotes. A continuación se componen las simetrías respecto de los lados en que se han ido produciendo los rebotes, y si h es la transformación así obtenida, h^2 debe ser una traslación. Se procede entonces a estudiar la trayectoria de un elemento (P, θ) que tenga orientación θ igual a la de la traslación hallada previamente.

Si el billar corresponde a un polígono racional Q cuyo grupo diedral asociado es suficientemente sencillo, es posible que la condición necesaria demostrada en el punto anterior sea también suficiente. Entre los billares que suponemos pueden estar en esta situación están los tres casos estudiados en 3.4.2, es decir, los triángulos racionales, los polígonos regulares, y más en general polígonos racionales de lados de igual longitud. (Tal vez con alguna restricción adicional respecto al grupo diedral D asociado a Q o respecto a la traslación elegida).

3.6 DENSIDAD DE LAS ORBITAS EN BILLARES POLIGONALES RACIONALES

En este apartado suponemos que Q es un billar poligonal racional, no necesariamente convexo, y cuyo borde ∂Q puede ser no conexo (billar con obstáculos). La hipótesis de racionalidad permite la construcción de un espacio de fases simplificado, sobre el que se puede aplicar la teoría de isometrías locales.

3.6.1 Espacio de fases

Sabemos por i) de 3.2.3. que si (Q, r) es un billar racional, el número máximo de orientaciones distintas que pueden adoptar los elementos de trayectoria $r^n(P, \theta)$ contruidos a partir de un elemento inicial (P, θ) , es un número finito, igual al orden del grupo diedral asociado al billar. Podemos construir un espacio de fases adaptado a la orientación particular θ del elemento de trayectoria inicial, que llamaremos S_θ , y que consiste en una unión

finita de intervalos de diámetro positivo. Consideremos para ello el conjunto $A_\theta = \{ \theta_i \}$, $1 \leq i \leq p$ de todas las posibles orientaciones que pueden adoptar los $\tau^n(P, \theta)$, y que coinciden con las distintas orientaciones de la órbita de (P, θ) por el grupo diedral D asociado al billar. Consideremos también el conjunto de los lados de ∂Q , $L = \{ l_j \}$, $1 \leq j \leq m$, y sea:

$$S_\theta = L \times A_\theta = \{ C_{ij} \}, \quad 1 \leq i \leq p; \quad 1 \leq j \leq m$$

Consideremos cada C_{ij} es una copia de l_j en la que, si β_{ij} es el ángulo que forma la orientación θ_i con el lado l_j y d representa la métrica usual en el plano, definimos una métrica d_{ij} :

$$d_{ij}(P_1, P_2) = d(P_1, P_2) \cdot |\sin \beta_{ij}|, \quad P_1, P_2 \in C_{ij}$$

Consideramos a cada C_{ij} dotado de la topología derivada de esta métrica, y en S_θ consideramos la topología de la suma topológica de sus componentes C_{ij} .

El espacio de fases S_θ así obtenido, no es sino un subconjunto del espacio de fases Q del billar, verificándose que $\tau(S_\theta) \subseteq S_\theta$, de manera que toda la órbita de (P, θ) está incluida en S_θ , y el estudio de dicha órbita se puede realizar sobre el sistema dinámico (S_θ, τ) (aquí τ es en realidad la restricción de τ a S_θ , que denotamos de la misma forma para simplificar).

3.6.2 Isometría local inducida

Probaremos a continuación que:

i) $\tau : S_\theta \longrightarrow S_\theta$ es una reordenación por intervalos generalizada.

Señalemos en primer lugar que el hecho de que dos elementos de trayectoria que parten con la misma orientación de un mismo lado — es decir, que pertenecen a un mismo C_{ij} — e inciden también en un mismo lado sin pasar por un vértice, rebotan de forma que la anchura de la banda que determinan permanece invariante. Pero la anchura de tal banda, es precisamente la distancia d_{ij} de los dos elementos. Esto prueba que todo elemento de trayectoria que no pasa por un vértice, está incluido en un intervalo de diámetro positivo que se transforma por τ isométricamente. Los elementos de S_θ que no se hallan en esta situación, es decir, aquellas semirrectas con origen en ∂Q que pasan por los vértices, no exceden de uno por cada vértice y orientación, y forman por tanto un conjunto finito, $\{x_i\}$, $1 \leq i \leq s$. Dichos elementos dividen cada C_{ij} en un número finito de componentes isométricas, y la restricción de τ a las componentes isométricas, es inyectiva. Por tanto τ es una reordenación de intervalos generalizada en S_θ . ■

Recordemos ahora que la definición de τ en sus puntos de discontinuidad como la identidad, tenía por motivo la conveniencia técnica de que τ fuera una aplicación, y no suponía una

aproximación real al modelo físico del billar. El resultado ii) de 2.7.1, permite asegurar que modificando adecuadamente la definición primitiva de τ en sus puntos de discontinuidad y quitando eventualmente algunos puntos de los extremos de los C_{ij} , puede conseguirse la continuidad lateral en todos los puntos, resultado que si representa una aproximación al modelo físico. En adelante supondremos que τ y S_0 son los así obtenidos.

El siguiente ejemplo, prueba que en general g no es orientable:

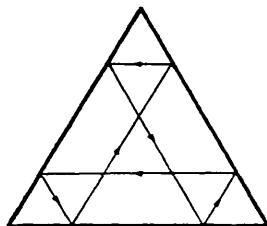


fig 3.8

Tomamos un triángulo equilátero de base horizontal, tomamos su base como origen de referencia de ángulos, y elegimos una orientación inicial $\theta = \pi/3$. Si l es la base del triángulo, uno de los segmentos de S_0 es (l, θ) , y además (ver fig. 3.8) τ^3 restringida a (l, θ) es una involución sobre este segmento, de manera que independientemente de como se defina el orden en los C_{ij} de S_0 , τ no puede ser creciente en cada C_{ij} , y por tanto no es orientable.

3.6.3. Estructura de una órbita cualquiera.

Aplicando los teoremas de recurrencia y estructura para reordenaciones por intervalos cualesquiera (ver 2.7.2) se obtiene:

i) En un billar poligonal racional, la órbita de cualquier elemento (P, θ) es densa en una unión finita de intervalos del espacio de fases S_0 , que en el caso de órbitas periódicas se reducen a puntos. Traducido este resultado en términos de la trayectoria descrita por un punto material rebotando en una mesa de billar poligonal racional, significa que la trayectoria es densa en una región formada por una unión finita de bandas, que en el caso de órbitas periódicas, se reducen a segmentos, y en el caso no periódico son de anchura positiva.

3.6.4. Caso genérico.

Finalmente probaremos que en el caso genérico la trayectoria cubre densamente toda la mesa, caracterizando los casos singulares. Supondremos aquí que el billar es racional, y que el interior del polígono Q consta de una sola componente conexa (es fácil, si se prescinde de esta hipótesis, construir un obstáculo que aisle cierta parte de la mesa).

Comenzamos probando el siguiente resultado:

i) Supongamos que θ es tal que en (Q, τ) , la trayectoria de ningún $x \in S_\theta$ une dos vértices. Entonces, τ es una reordenación por intervalos en S_θ , regular y acíclica.

Supongamos que τ no es regular. Existen puntos $x, y \in X(\tau)$ tales que para cierto $n \in \mathbb{N}$, $\tau^n(x) = y$. Pero los puntos de cambio de τ , sabemos que son elementos de trayectoria que pasan por los vértices, por lo que $\tau(x)$ y $\tau(y)$ tienen su origen en vértices, y la trayectoria de $\tau(x)$ une estos dos vértices.

Supongamos ahora una trayectoria periódica en Q , correspondiente a algún elemento de S_θ . Por la continuidad lateral de τ , esta trayectoria se puede incluir en una banda formada enteramente por trayectorias periódicas, paralelas entre sí. Dicha banda, se puede ensanchar hasta llegar a algún vértice A . Consideramos ahora la trayectoria que parte de A con la misma orientación que el rebote de la banda en el lado del billar en que está A . Esta trayectoria será frontera de la banda, hasta que encuentre un nuevo vértice y se separe de ella, o bien hasta que regrese a A . En ambos casos es cierta la proposición. ■

ii) El conjunto de orientaciones θ que verifican la anterior propiedad, es numerable y denso en $[0, 2\pi)$.

Es numerable, por ser numerable el conjunto de las trayectorias que unen dos vértices, ya que si usando el método del desplazamiento de la mesa (ver 3.3.), prolongamos en línea recta una trayectoria de este tipo, mientras que vamos reflejando la mesa en los lados en que la prolongación de la trayectoria incide, se llegará a obtener un segmento de recta que une un vértice de la mesa inicial Q a un vértice de la n -ésima mesa, $Q_n = g(Q)$, donde g es cierta transformación del grupo R (ver 3.2.1.). Como dicho grupo es numerable, los posibles $g(Q)$ también lo son, y el conjunto de segmentos que unen vértices de $g(Q)$ a vértices de Q , también lo es. Para probar que es denso, tomamos un elemento de trayectoria cualquiera (P, θ) de (Q, τ) , donde P es un vértice, y tomamos un cono $C(P, \beta)$ de vértice P y ángulo β arbitrariamente pequeño que contiene a (P, θ) en su interior. Prolongando suficientemente el elemento de trayectoria por el método de la mesa móvil, se puede conseguir que la mesa Q_n esté incluida en $C(P, \beta)$, y uniendo P con cualquiera de los vértices de Q_n se encuentra una trayectoria que une dos vértices y tiene orientación más próxima a θ que el ángulo β . ■

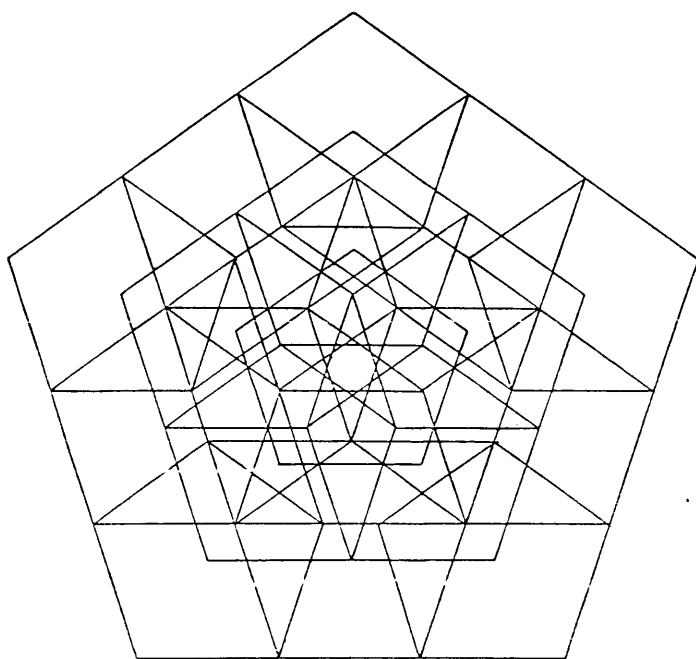
Llamaremos a las orientaciones de este tipo orientaciones singulares, mientras que el resto de las orientaciones serán orientaciones genéricas. Finalmente se prueba:

iii) Si θ es una orientación genérica, la trayectoria de cualquier elemento de S_θ es densa en el polígono Q .

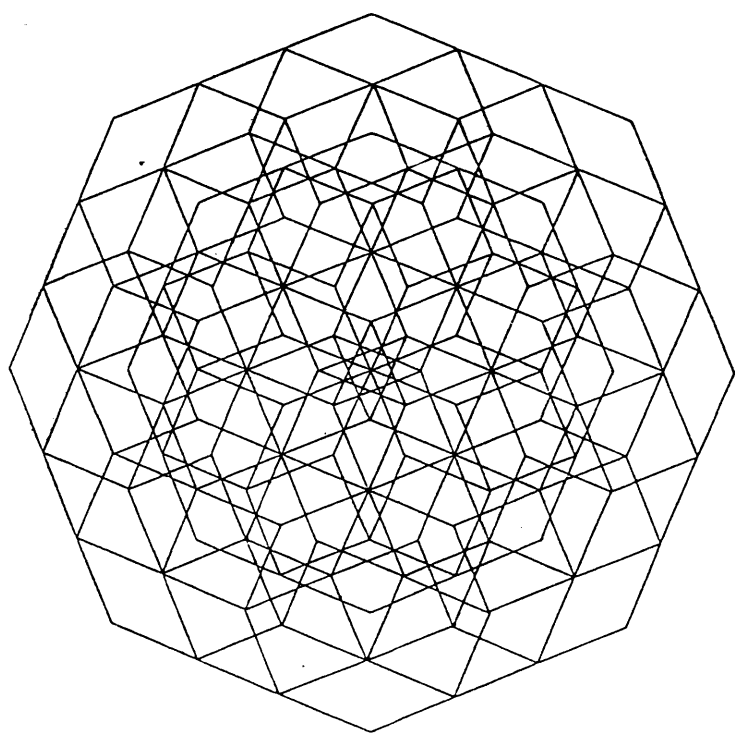
Teniendo en cuenta que ahora τ es una reordenación por intervalos en S_θ , regular y acíclica, y que las componentes conexas de S_θ son de la forma (I, α) , del teorema vi) de 2.6.5 se deduce que si en una trayectoria hay un (P, α) , con $P \in I$, la trayectoria es densa en (I, α) , y también que si una trayectoria incide en un lado I con una orientación α , una banda de orientación α en la que los elementos de orientación α son densos, cubre el lado I . Si una trayectoria O no es densa en Q , y cierta región G de la mesa no es atravesada por ningún elemento de trayectoria, dicha región no puede estar limitada exclusivamente por lados de ∂Q , porque G coincidiría con todo Q , o bien Q no sería conexo. G debe estar limitada por algún elemento de trayectoria ξ y por lados de ∂Q . Consideremos el elemento de trayectoria vertical, y que G está a su derecha. A la derecha de ξ hay toda una banda vertical sin elementos verticales de la trayectoria O . Tomemos un punto $P \in \text{Fr}(G) \cap \xi$, y tracemos una vertical partiendo de P hacia arriba hasta que corte un lado I de ∂Q en un punto A . A la derecha de A , hay en I puntos no visitados por elementos verticales de O . Pero si A fuese un punto interior de I , todo I estaría incluido en una banda en que los elementos verticales de O serían densos, luego A debe ser el extremo derecho de I , es decir, un vértice. Realizando el mismo razonamiento con la vertical desde P hacia abajo, obtenemos que z une dos vértices, y θ es singular. ■

ANEXO DE GRAFICOS DEL

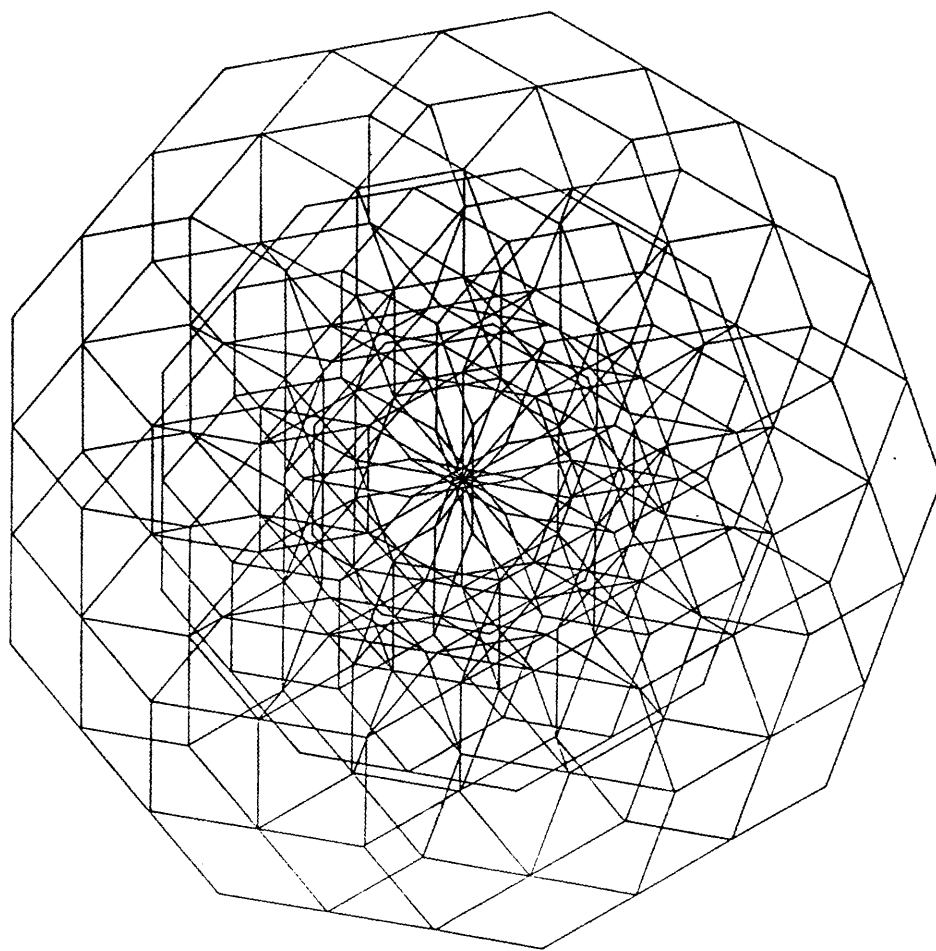
CAPITULO III



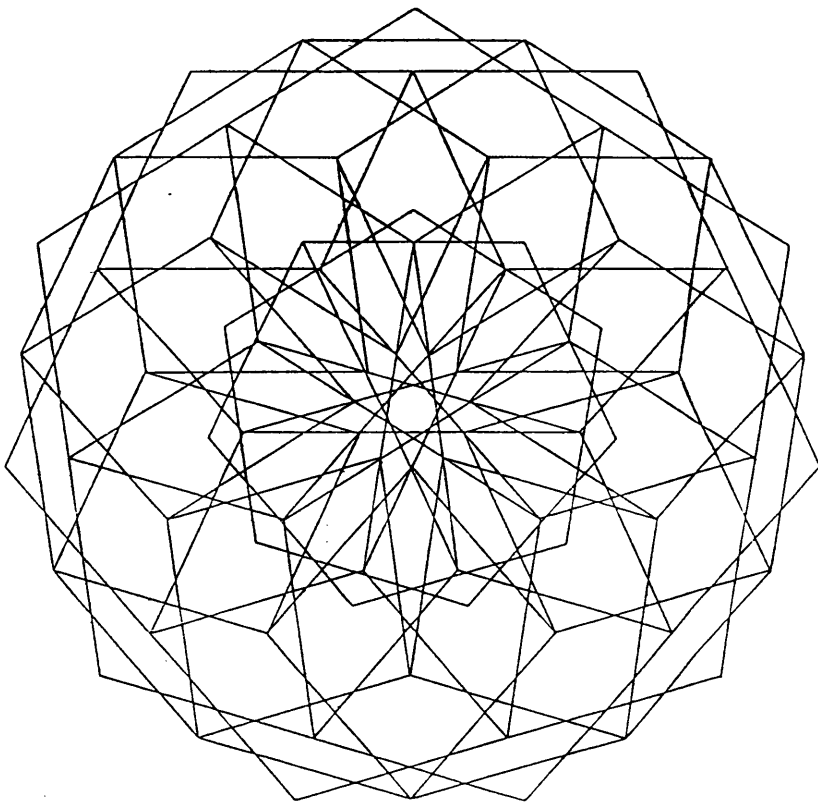
Diseño obtenido mediante el anillo A_5 (descripción en 3.4.3).



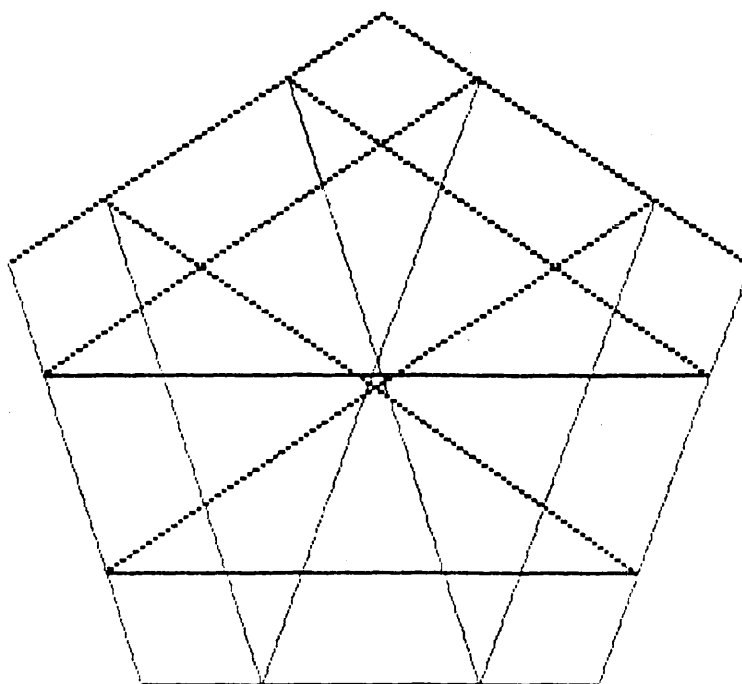
Diseño obtenido mediante el anillo Ag.



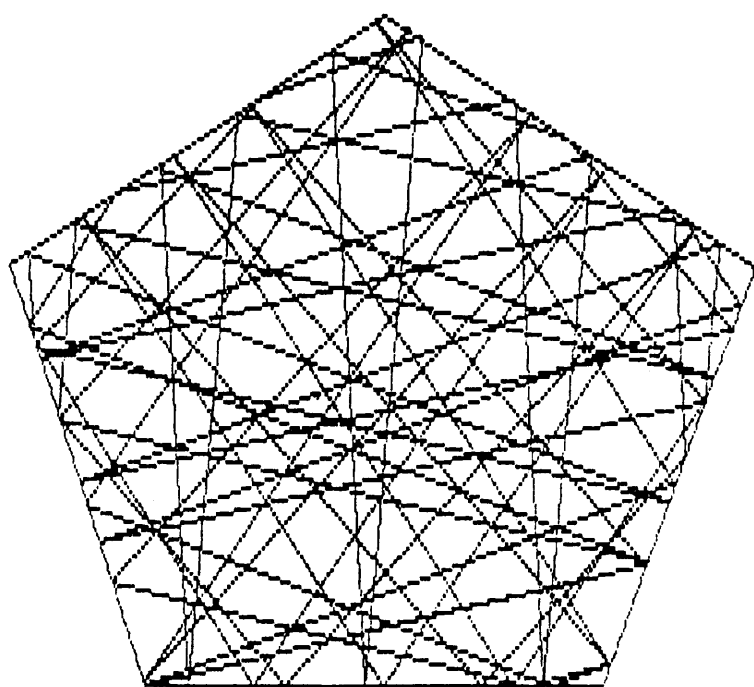
Diseño obtenido mediante el anillo Ag.



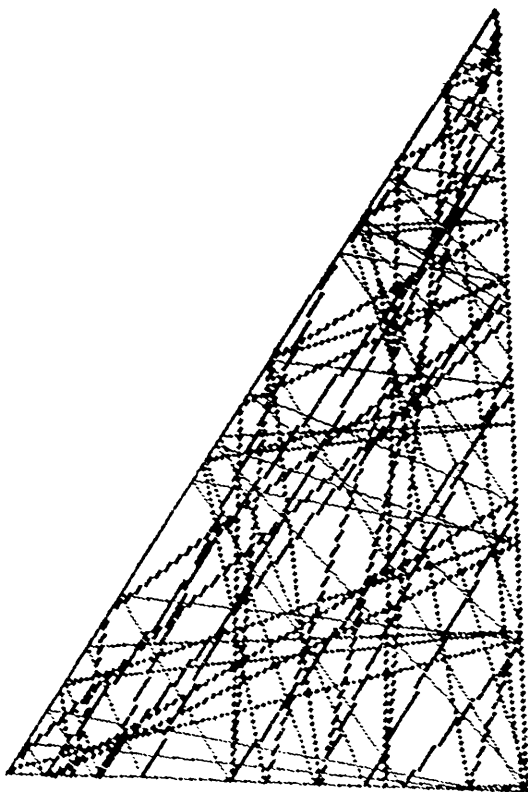
Diseño obtenido mediante el anillo A_{11} (en este caso se unen vértices alternos).



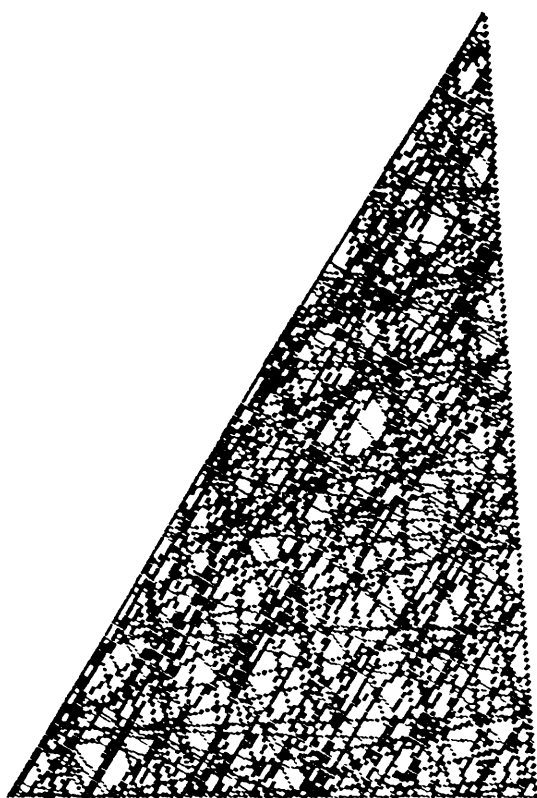
Orbita periódica en un pentágono regular.
Número de rebotes: 10



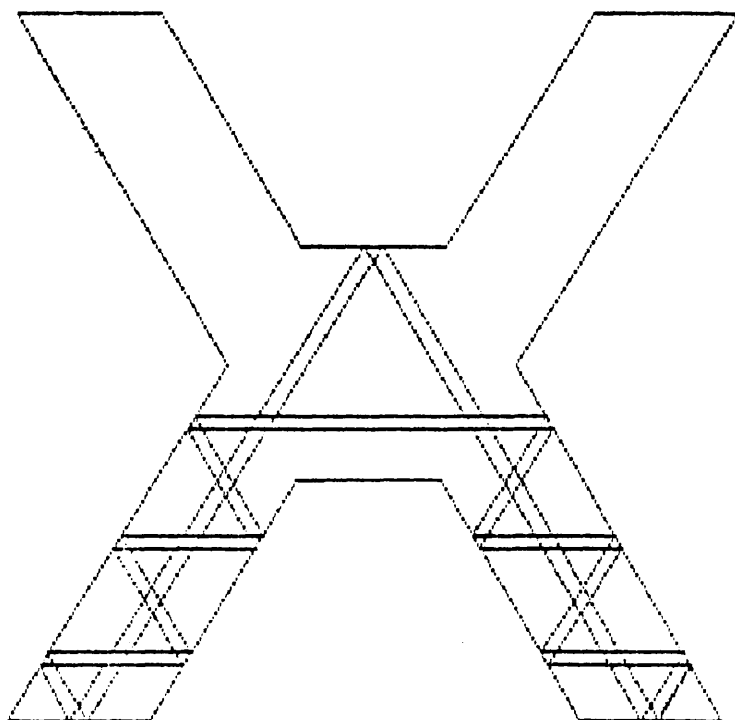
Orbita periódica en un pentágono regular
Número de rebotes: 58



Orbita periódica en billar triangular (ángulos $\pi/3$ y $\pi/5$).
 Número de rebotes: 90



Orbita periódica obtenida en triángulo (ángulos $\pi/3$ y $\pi/5$)
Número de rebotes : 208



Orbita periódica obtenida en un billar de tipo c) descrito en 3.4.2. En este caso los lados no son iguales pero son de longitud múltiplo del más pequeño, por lo que es equivalente al citado caso. Número de rebotes: 26



Orbita periodica obtenida en el mismo billar que el anterior.
Número de rebotes: 846

LISTADO DEL PROGRAMA USADO EN LOS EJEMPLOS ANTERIORES.

LENGUAJE DE PROGRAMACION : Quick Basic 3.0

```
'EN ESTE PROGRAMA EL ORIGEN DE COORDENADAS ESTA EN EL
'ANGULO INFERIOR
'IZQUIERDO, Y CABEN 200x200 PUNTOS
DEFDBL A-H,K-M,O-Z
DEFINT I-J,N
INPUT "¿COMPROBACION DE TRASLACIONES, SI O NO?";C$
IF C$="SI" THEN ID=0 ELSE ID=1
'ENTRADA DE DATOS
'DATOS DE LA MESA
DIM V(2,50) 'VERTICES DEL PERIMETRO
PI=3.141592653589793
DIM IT(1000)
INPUT "NUMERO TOTAL DE VERTICES";N
'BASE DEL POLIGONO
INPUT "COORD. DE LOS DOS VERT. DE LA BASE( 1º EL DERECHO)";_
;V(1,1),V(2,1),V(1,2),V(2,2)
PRINT "ANGULOS EN ESTOS VERTICES EN FRACCION DE  $\pi$ ,"
INPUT "DOS NUMERADORES Y DOS DENOMINADORES";_
;BETN,BETD,ALFN,ALFD
BETA=BETN/BETD
ALFA=ALFN/ALFD
ALFA=ALFA*PI:BETA=BETA*PI
ALFSUM=ALFA
FOR I= 3 TO N-1
PRINT"LONGITUD DEL LADO"; I-1;"?"
INPUT LONG
ANG = ALFA
GOSUB 4000
GOSUB 5000
V(1,I)=V(1,I-1) + COSENO*LONG
V(2,I)=V(2,I-1) + SENOS*LONG
PRINT "COORDENADAS DEL NUEVO VERTICE";V(1,I),V(2,I)
INPUT "ANGULO EN ESTE VERTICE ";ANVN,ANVD
ANV=ANVN/ANVD
ANV=ANV*PI
ALFSUM = ALFSUM+ANV
ALFA=ALFSUM-(I-2)*PI
NEXT I
'CALCULO DEL ULTIMO VERTICE
ANG=ALFA
GOSUB 4000
GOSUB 5000
IF COSENO = 0 THEN MI = 1.D20 ELSE MI=SENO/COSENO
ANG = PI-BETA
GOSUB 4000
GOSUB 5000
IF COSENO = 0 THEN M =1.D20 ELSE M = SENO/COSENO
A1=M :B1=-1 :C1=-M*V(1,1)+V(2,1)
A2=MI :B2=-1 :C2=-MI*V(1,N-1)+V(2,N-1)
GOSUB 1000
```

```

V(1,N)=XN : V(2,N)=YN
PRINT"COORDENADAS DEL ULTIMO VERTICE";V(1,N),V(2,N)
'ENTRADA DE DATOS TRAYECTORIA
'MATRICES PARA PUNTOS DE TRAYECTORIA EN LOS BORDES DE LA MESA
DIM X(1000),Y(1000)
'CALCULO DE COEFICIENTES DE LAS ECUACIONES DE LOS LADOS
DIM L(3,N) 'MATRIZ PARA LOS COEFICIENTES
FOR I = 1 TO N
IF I=N THEN J=1 ELSE J=I+1
X1=V(1,I):Y1=V(2,I)
X2=V(1,J):Y2=V(2,J)
GOSUB 2000
L(1,I)=A:L(2,I)=B:L(3,I)=C
NEXT I
'CONTADOR PARA REPETIR PRUEBAS DE TRASLACIONES
IR=0
7 INPUT"UN PAR DE COORDENADAS Y UN ANGULO INICIALES ";_
X(0),Y(0),ALFN,ALFD
ALFA=ALFN/ALFD
8 ALFA=ALFA*PI
IN=0
'IMPRESION DE LA MESA
SCREEN 2
FOR I = 1 TO N
IF I=N THEN J=1 ELSE J= I+1
LINE(3.2*V(1,I),200-V(2,I)*1.275)-(3.2*V(1,J),200-V(2,J)*1.275)
NEXT I
'CALCULO DE LA TRAYECTORIA
'CALCULO DE LOS COEFICIENTES EN EL PRIMER TRAMO
'MATRICES DE COEFICIENTES
DIM CA(1000),CB(1000),CC(1000)
'MATRICES AUXILIARES PARA LA SELECCION DEL LADO EN QUE REBOTA
DIM IA(50),XA(50),YA(50)
X1=X(0):Y1=Y(0)
IF ID = 1 AND IR=1 THEN GOTO 80
ANG=ALFA
GOSUB 4000
GOSUB 5000
X2=X(0)-COSENO:Y2=Y(0)-SENO
XC=X2:YC=Y2 'VARIABLES PARA DISTINGUIR TRASLACIONES
'Y GIROS DE 180
80 GOSUB 2000
CA(1)=A:CB(1)=B:CC(1)=C
'VARIABLES PARA CUANDO SE REBASAN LOS 1000 REBOTES
GA=A:GB=B:GC=C
IC=1
'CALCULO DE LOS CORTES CON LOS LADOS
1 IF IC > 1000 GOTO 201
A1=CA(IC):B1=CB(IC):C1=CC(IC)
201 A1=GA:B1=GB:C1=GC
FOR IK=1 TO N
IF IK=IN GO TO 2
IF IK=N THEN J=1 ELSE J = IK+1
A2=L(1,IK):B2=L(2,IK):C2=L(3,IK)

```

```

GOSUB 1000
IF(V(1,IK)-XN)*(V(1,J)-XN)>1.D-8 OR (V(2,IK)-YN)*_
(V(2,J)-YN)>1.D-8 GOTO 2
IF(X2-X1)*(XN-X1)>0 OR (Y2-Y1)*(YN-Y1)>0 GOTO 2
IA(1K)=1:XA(1K)=XN:YA(1K)=YN
GOTO 3
2 IA(1K)=0
3 NEXT IK
'SELECCION DEL CORTE MAS PROXIMO A (X1,Y1)
S=10000
FOR I =1 TO N
IF IA(I)=0 GOTO 4
D= SQR((X1-XA(I))^2 +(Y1-YA(I))^2)
IF D<S THEN S=D : IN=I
NEXT I
IF IC>1000 GOTO 104
IT(IC)=IN
104 'IMPRESION DEL ELEMENTO DE TRAYECTORIA
LINE(3.2*X1,200-Y1*1.275)-(3.2*XA(IN),200-YA(IN)*1.275)
IF TR = 1 THEN TR = 0 : IR=1 :GOTO 101
'CALCULO DEL NUEVO ELEMENTO DE TRAYECTORIA
XX=X1:YY=Y1
A1=L(1,IN):B1=L(2,IN):C1=L(3,IN)
GOSUB 3000
X2=XS:Y2=YS
X1=XA(IN):Y1=YA(IN)
IF IC>1000 GOTO 203
X(IC)=X1:Y(IC)=Y1
203 GOSUB 2000
IC=IC+1
IF IC>1000 GOTO 202
CA(IC)=A:CB(IC)=B:CC(IC)=C
202 GA=A:GB=B:GC=C
IF ID = 0 AND C$ = "SI" THEN GOTO 204
IF ABS(GA*X(0) +GB*Y(0)+GC)>1.D-6 GOTO 204
IF SGN(X2-X1)=SGN(XC-X(0)) AND SGN(Y2-Y1)=SGN(YC-Y(0))_
THEN GOTO 504 ELSE GOTO 204
504 IF ABS(CB(1)) >1.D-10 AND ABS(GB) > 1.D-8 THEN GOTO 500
IF ABS(CB(1)) <1.D-10 AND ABS(GB) < 1.D-8 THEN GOTO 501_
ELSE GOTO 204
500 IF ABS(GA/GB-CA(1)/CB(1))>1.D-6 THEN GOTO 204
501 PRINT "TRAYECTORIA PERIODICA. OTRO CASO ¿SI O NO?"
TR=1
GO TO 1
204 'COMPROBACION DE TRASLACIONES
IF ID = 1 GOTO 30
IF (IC-1) MOD 2 = 1 GOTO 30
IF ABS(CA(1))<1.D-8 OR ABS(CB(1))<1.D-8 THEN GOTO 20
IF ABS(CA(IC)/CA(1)-CB(IC)/CB(1))>.00001 THEN_
GOTO 30 ELSE GOTO 40
20 IF (ABS(CA(1))< 1.D-8 AND ABS(CA(IC))<1.D-8)_
OR (ABS(CB(1))<1.D-8 AND ABS(CB(IC))<1.D-8) THEN GOTO 40
30 IF IC = 1000 AND ID=1 AND IR=0 THEN GOTO 42
IF IC = 5000 THEN PRINT"5000 REBOTES, OTRO CASO,¿SI O NO?"_

```

```

:GOTO 101
A$=INKEY$
IF A$="" GOTO 1
CLS
IF ID=0 OR IR=0 THEN GOTO 7 ELSE GOTO 70
42 PRINT "1000 REBOTES. ¿OTRO CASO?, SI O NO"
101 INPUT D$
IF D$="NO" THEN END
CLS
IF D$="SI" THEN IF IR=1 THEN PRINT_
"¿CON LA MISMA ORIENTACION?, SI O NO" ELSE GOTO 7
INPUT F$
IF F$="SI" GOTO 70 ELSE IR=0:ID=0 :GOTO 7
40 IF SGN(X2-X1)=SGN(XC-X(0)) AND SGN(Y2-Y1)=SGN(YC-Y(0))_
THEN GOTO 41 ELSE GOTO 30
41 PRINT "TRASLACION EN EL REBOTE"; IC-1 ; "¿CONTINUAR O PROBAR?"
50 B$=INKEY$
IF B$="" GO TO 50
IF B$="C" GOTO 30
IF B$="P" GO TO 60
PRINT "MENSAJE ERRONEO. VUELVA A INTENTARLO"
GOTO 50
60 'COMPROBACION DE SI UNA TRASLACION DA UNA ORBITA PERIODICA
SCREEN 0
CLS
'CALCULO DE LA TRASLACION
XS=X(0):YS=Y(0)
FOR JJ = 1 TO 2
FOR I= 1 TO IC-1
A1=L(1,IT(I))
B1=L(2,IT(I))
C1=L(3,IT(I))
XX=XS:YY=YS
GOSUB 3000
NEXT I
NEXT JJ
XT=X(0)-XS
YT=Y(0)-YS
70 INPUT "UN PAR DE COORDENADAS INICIALES";X(0),Y(0)
'CONTADOR PARA QUE NO SE BUSQUEN TRASLACIONES CUANDO
'SE ESTA COMPROBANDO
'LA PERIODICIDAD DE UNA DE ELLAS
ID=1
'CONTADOR PARA REPETIR LA PRUEBA CON OTRO PUNTO
'INICIAL PERO LA MISMA
'ORIENTACION DE TRASLACION
IR=1
X2=X(0)-XT
Y2=Y(0)-YT
XC=X2:YC=Y2
GOTO 8

```

```

'SUBROUTINAS
1000 D=A1*B2-A2*B1
    IF ABS(D)<1.D-20 THEN XN=1.D10:YN=1.D10:RETURN
    DX=-C1*B2+C2*B1
    DY= -A1*C2+A2*C1
    XN=DX/D
    YN=DY/D
    RETURN
2000 A=Y2-Y1
    B=X1-X2
    C= -A*X1-B*Y1
    RETURN
3000 A2=-B1
    B2=A1
    C2=B1*XX-A1*YY
    GOSUB 1000
    XD= XN-XX
    YD= YN-YY
    XS= XN+XD
    YS= YN+YD
    RETURN
4000 SE=ANG
    SENO=ANG
    FOR IW= 1 TO 20
    SE =-1* SE*ANG*ANG/((2*IW*(2*IW+1))
    SENO=SENO+SE
    NEXT IW
    RETURN
5000 CO = 1
    COSENO = 1
    FOR IW = 1 TO 20
    CO = -1*CO*ANG*ANG/((2*IW-1)*2*IW)
    COSENO = COSENO+CO
    NEXT IW
    RETURN

```

CAPITULO IV

ISOMETRIAS LOCALES BIYECTIVAS EN \mathbb{R}^2

4.0. INTRODUCCION.

Este capítulo trata sobre isometrías locales biyectivas en R^2 , con aplicaciones a billares poliédricos racionales.

En el primer apartado se introduce el concepto de isometría local biyectiva en R^2 sobre la base del ejemplo que brindan los billares poliédricos racionales, y se muestran también un par de ejemplos abstractos cuya dinámica puede analizarse completamente gracias a su especial sencillez, enumerándose algunos tipos particulares de i.l.b. en R^2 que luego son tratados en el apartado tercero.

En el segundo apartado se obtienen algunas propiedades generales de las i.l.b. en R^2 con consecuencias en billares poliédricos racionales, particularmente la recurrencia de estos billares, la caracterización de la familia de orientaciones de órbitas periódicas como excepcionales, y en el caso de orientaciones genéricas, la demostración de que las órbitas se aproximan arbitrariamente a las aristas del poliedro.

En el tercer apartado se analizan i.l.b. en R^2 obtenidas por producto cartesiano de i.l.b. en R . Se analiza la dinámica en caso de que alguno de los factores sea periódico, y se obtiene una clasificación de cierto tipo de automorfismos de un toro, según las propiedades de sus órbitas. Se obtiene una condición suficiente para que un producto de reordenaciones por intervalos sea casi-minimal, y se obtienen algunas propiedades interesantes para el caso $f \times f$, caso que presenta el interés de tener aplicaciones en el estudio de billares poligonales racionales con dos bolas.

Finalmente en el último apartado se hace una aproximación al problema de billares poligonales no racionales, mostrándose que conduce a un tipo de modelo dinámico próximo en muchos aspectos al de las isometrías locales biyectivas.

4.1 ISOMETRIAS LOCALES BIYECTIVAS EN R^2 .

4.1.1. Billares poliédricos racionales

Sea ahora Q una región acotada, abierta y conexa de R^3 limitada por una superficie poliédrica ∂Q , y supongamos un punto material en movimiento rectilíneo perpetuo en el interior de Q , rebotando en ∂Q según las leyes de la reflexión elástica. Este modelo dinámico es un billar poliédrico. Si consideramos, como en 3.2.2, el conjunto de las simetrías de R^3 respecto a planos paralelos a las caras de ∂Q , pasando todos ellos por un punto dado O , y llamamos D al grupo generado por estas simetrías, si este grupo es finito, según razonamientos similares a los de i) de 3.2.3, se puede establecer que las distintas orientaciones de los elementos una trayectoria en el billar poliédrico, pertenecen a un conjunto

finito . En este caso , diremos que el billar poliédrico es racional.

Describiremos las orientaciones de los elementos de trayectoria mediante el conjunto U de los vectores de R^3 que tienen módulo unidad. Un elemento de trayectoria en el billar puede caracterizarse como (P, θ) donde $P \in \partial Q$ y $\theta \in U$. Si tomamos un elemento de trayectoria inicial (P, θ) y el billar es racional, llamaremos A_θ al subconjunto finito de U de todas las posibles orientaciones que pueden alcanzar los elementos de la trayectoria de (P, θ) . Si L es el conjunto finito de caras de ∂Q , se puede definir un espacio de fases $S_\theta = L \times A_\theta$ y una aplicación $\tau: S_\theta \rightarrow S_\theta$ que describen la dinámica del billar poliédrico, para trayectorias de orientación de A_θ . La aplicación τ es la restricción a S_θ de la aplicación de reflexión, cuya definición es la dada en 3.1.3, sustituyendo las simetrías axiales por simetrías respecto a los planos que contienen las caras del poliedro si (P, θ) es tal que incide en el interior de una cara, mientras que cuando incide en una arista, puede ser definida como la identidad. La aplicación τ es discontinua en las preimágenes de las aristas de Q , que formarán curvas poligonales sobre las caras de ∂Q . Cada $(c, \theta) \in S_\theta$, donde c es una cara de ∂Q , — llamaremos a estos elementos caras de S_θ — se verá dividido por una poligonal en polígonos en cuyo interior τ es una aplicación continua — llamaremos a estos polígonos componentes isométricas de τ —, y según probaremos, isométrica respecto a una métrica adecuada.

La simetría respecto a un plano en el espacio es una isometría. Si dos rectas r_1 y r_2 son paralelas, sus simétricas respecto a un plano también lo son y la distancia entre r_1 y r_2 coincide con la de sus simétricas. Sean $x_1 = (P_1, \theta)$ y $x_2 = (P_2, \theta)$ elementos de trayectoria de $(c, \theta) \in S_\theta$. Si representamos por $[x_1]$ y $[x_2]$ respectivamente la recta que pasa por P_1 y tiene vector de dirección θ y que pasa por P_2 y tiene el mismo vector de dirección, definiendo la distancia entre x_1 y x_2 como la distancia entre las rectas $[x_1]$ y $[x_2]$ obtenemos una métrica en (c, θ) . Haciendo lo propio en cada cara de S_θ , obtenemos una métrica en S_θ tal que la restricción de τ a cada una de las componentes isométricas de cada cara, es una isometría. Este proceso es equivalente a la identificación de cada cara (c, θ) con la proyección de c sobre un plano perpendicular a la orientación θ . De esta forma, se puede definir en cada (c, θ) una medida μ , igual a la medida de Lebesgue en la cara c multiplicada por un factor de normalización, igual al coseno del ángulo formado por la cara y el plano perpendicular a θ , y hacer así de S_θ un espacio de medida tal que la aplicación de reflexión τ conserva la medida.

Vemos así que el estudio de los billares poliédricos, se resuelve en el marco de las isometrías locales en R^2 , al que dedicamos el resto del apartado.

4.1.2. Definición de isometrías locales biyectivas en R^2 .

El punto anterior sugiere la siguiente *definición de isometría local biyectiva en R^2* :

Sea f una aplicación, $f: S \longrightarrow S$, $S \subset R^2$. Decimos que f es una *isometría local biyectiva en R^2* si:

a) S es una unión finita de conjuntos acotados homeomorfos a intervalos de R^2 (producto de intervalos abiertos, cerrados o semicerrados de R).

b) El conjunto G de puntos de discontinuidad de f es unión finita de curvas continuas rectificables, y si $x \in S - G$, hay un S -entorno de x , $V(x) \subset S - G$ tal que $f|V(x)$ es isometría.

c) La restricción de f a $S - G$ es inyectiva.

Esta definición es verificada por $\tau: S_0 \longrightarrow S_0$, por lo que condiciones suplementarias que impliquen la transitividad, minimalidad, ergodicidad etc. de las isometrías locales biyectivas, se traducen en las correspondientes propiedades para el billar poliédrico.

Exponemos a continuación dos ejemplos en los que puede además clasificarse la dinámica del sistema.

4.1.3. Dos casos particulares.

Supongamos ahora que S es un intervalo de R^2 como el de la fig. 4.1. En dicha figura, las rayas de trazo grueso corresponden a la curva G de puntos de discontinuidad, y las componentes isométricas

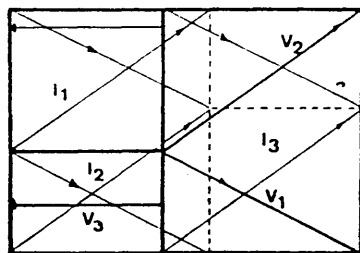


fig. 4.1

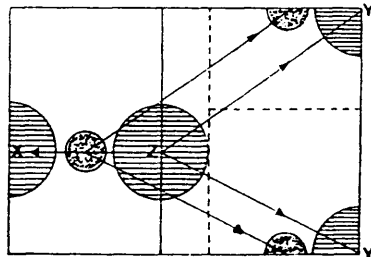


fig 4.2

son los tres intervalos determinados por ellas, I_1, I_2, I_3 . Se define la i.l.b. $f: S \rightarrow S$ trasladando cada rectángulo respectivamente por los vectores v_1, v_2 y v_3 señalados en la figura. Obsérvese que si consideramos todos los intervalos cerrados por su borde inferior y por su borde izquierdo y abiertos en los otros dos bordes, f es no solamente inyectiva en su restricción a $S - G$, sino que es biyectiva en todo S y hay continuidad "lateral" en los puntos de G , en el sentido de que si $x \in G$, podemos encontrar un intervalo I_x tal que $x \in I_x$ y $f|_{I_x}$ es isometría. Naturalmente I_x no es entorno de x , ya que x estará en la frontera de I_x .

Completamos ahora S con el borde superior y borde derecho, y procedamos a identificar cada punto del borde inferior de S con el punto del borde superior que está en su vertical. La figura 4.2 muestra como un disco centrado en un punto de la línea horizontal de G que se transformaría en dos semidiscos disjuntos por f , se transforma ahora, teniendo en cuenta las identificaciones, en un disco conexo. Teniendo en cuenta que ahora los bordes laterales son círculos, podemos identificarlos previa rotación, de manera que un punto arbitrario del borde derecho resulte identificado con uno arbitrario del borde izquierdo. Si esto lo hacemos de manera que el punto x resulte identificado con el y (ver fig. 4.2), entonces la transformación T inducida por f en S' es continua en todos los puntos de S' ($T = g \circ f \circ g^{-1}$, siendo g la proyección canónica de S en S'). Observar en la figura como un disco centrado en z se transforma en un disco del toro, teniendo en cuenta las identificaciones). Es fácil ver que T es una traslación en el toro S' cuya dinámica es bien conocida: Si la traslación en el toro es según un vector $v = (v_1, v_2)$ y v_1 y v_2 son racionalmente independientes, las órbitas bajo T son densas en S' , y si son racionalmente dependientes, son densas en una geodésica cerrada del toro. Esto significa que las órbitas bajo f de los elementos de S , son respectivamente densas en S y densas en la intersección de una colección finita de rectas de dirección v equidistantes entre sí con S . En la figura 4.3 mostramos un ejemplo obtenido con computador.

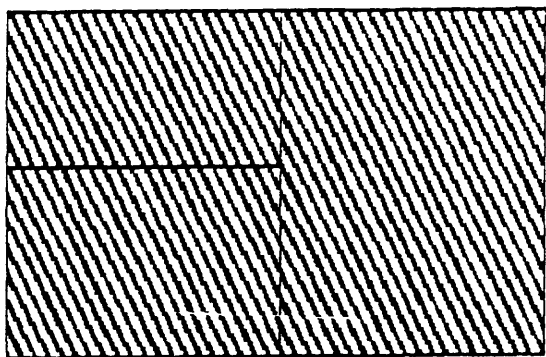


fig 4.3

Un ejemplo análogo al anterior, es la i.l.b. descrita en la figura 4.4. Aquí, tras identificar como antes los bordes inferior y superior de S , se procede a identificar los bordes laterales, de manera que $x = f(z)$ resulte identificado con y . En la figura se muestra como discos de centro en z y w se transforman en discos del toro, y la transformación inducida, es, como antes, una traslación en el toro.

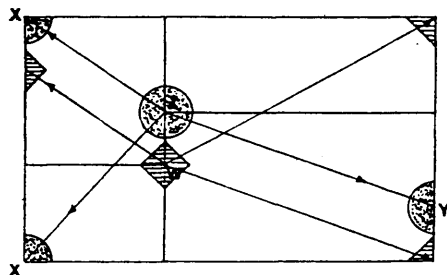


fig. 4.4

4.1.4. Algunos tipos de i.l.b. en R^2

Los dos ejemplos anteriores sugieren la siguiente definición de intercambio de intervalos en R^2 , como generalización de la de intercambio de intervalos en R (ver 2.1.1):

- i) $f: S \longrightarrow S$ es un *intercambio de intervalos en R^2* si:
 - a) f es biyectiva
 - b) S es un intervalo en R^2
 - c) Existe un recubrimiento $\{I_i\}, 1 \leq i \leq m$ de S donde los I_i son intervalos disjuntos, tales que $f|_{I_i}$ es una traslación.

ii) Si en la condición b) se admite que S sea unión de intervalos de R^2 , y en la condición c) permitimos que las restricciones $f|_{I_i}$ puedan ser el producto de una traslación por una simetría axial de eje vertical u horizontal, o por un giro de π radianes, obtenemos una generalización natural de la reordenación por intervalos en R , que llamaremos *reordenación por intervalos en R^2* (ver definición de r.i en R en 2.5.1).

iii) Un ejemplo de intercambio de intervalos en R^2 se puede obtener mediante el *producto de intercambios de intervalos*: Sean $T_1: S_1 \longrightarrow S_1$ y $T_2: S_2 \longrightarrow S_2$ dos intercambios de intervalos en R (no supondremos normalizada la longitud de S_1 y S_2). Consideramos entonces el intervalo $S = S_1 \times S_2$, y la aplicación

$f: S \rightarrow S$ definida por $f(x,y) = (T_1(x), T_2(y))$. Si suponemos que T_1 y T_2 son reordenaciones por intervalos en S_1 y S_2 , obtenemos un ejemplo de reordenación por intervalos en R^2 de tipo *producto de reordenaciones por intervalos*. Las componentes isométricas que forman el recubrimiento de S (ver punto 4.2.1), se obtienen como producto de componentes isométricas de T_1 por componentes isométricas de T_2 .

iv) Sea ahora un intercambio de intervalos $T: S_1 \rightarrow S_1$ y sean $\{J_i\}$, $1 \leq i \leq n$ sus componentes isométricas. Consideremos ahora una familia $\{T_i\}$, $1 \leq i \leq n$ de intercambios de intervalos, $T_i: S_2 \rightarrow S_2$, $1 \leq i \leq n$. Podemos definir un intercambio de intervalos $f: S = S_1 \times S_2 \rightarrow S$, que llamaremos *producto mixto de T por $\{T_i\}$* de la siguiente manera: Si $y \in J_i$, $f(x,y) = (T(x), T_i(y))$. Ejemplos de este tipo son los dos descritos en el punto anterior. En el primero, se trata del producto mixto de un intercambio de dos intervalos, por una familia formada por un intercambio de dos intervalos y la identidad. En el segundo ejemplo, la familia es de dos intercambios diferentes de dos intervalos. Las componentes isométricas de f se obtienen como producto de cada J_i por las componentes isométricas de T_i . Si en lugar de intercambios de intervalos tuviéramos reordenaciones por intervalos, obtendríamos análogamente un *producto mixto de reordenaciones por intervalos*. Es claro que el producto mixto engloba el caso anterior: Basta tomar una familia compuesta de intercambios iguales, y su producto mixto con otro cualquiera, coincide con el producto de este por uno de los de la familia.

4.2. PROPIEDADES GENERALES

4.2.1. Componentes isométricas.

Como en el caso de i.1.b. en R se dice que un conjunto $I \subset S$ es un *conjunto isométrico* si $f|_I$ es isometría, es *k-isométrico* si $f^k|_I$ es isometría si $1 \leq k \leq K$, *w-isométrico* si $f^w|_I$ es isometría y $i \in N$. La *componente isométrica* de $x \in S$ es el conjunto conexo maximal C_x , tal que $x \in C_x$, y tal que $f|_{C_x}$ es isometría. Una componente isométrica C_x no puede tener intersección no vacía con dos componentes conexas C_1 y C_2 de $S - G$, pues como en C_x no puede haber puntos de G , se obtendría C_x como unión de los C_x -cerrados disjuntos $C_x \cap C_1$, donde las C_i son las componentes conexas de $S - G$ y si dos de ellos son no vacíos, C_x es no conexa. Por otra parte, si $x \in S - G$, y es C'_x la componente conexa de x en $S - G$, $f|_{C'_x}$ debe ser isometría, ya que en otro caso en C'_x podríamos encontrar puntos de discontinuidad y sin embargo estos están en G . Las componentes isométricas deben ser por tanto componentes conexas de $S - G$, unidas tal vez a puntos de G que estén en su frontera.

4.2.2. Teorema de recurrencia.

De acuerdo con lo que precede, y también usando los argumentos de la demostración 2.2.1, es fácil obtener :

i) Si f es un i.l.b. en $S \subset \mathbb{R}^2$, f conserva la medida Lebesgue.

Como todas las hipótesis del teorema de recurrencia de Poincaré son satisfechas por (S, f, L_2) , si f es una i.l.b. en $S \subset \mathbb{R}^2$ y L_2 es la medida Lebesgue, obtenemos:

ii) Casi todos los puntos son recurrentes por f , es decir, para casi todos los puntos de S :

$$x \in \omega(x)$$

(Para las definiciones de los operadores ω y α ver 2.6.1)

4.2.3.

Nos proponemos aquí estudiar el conjunto $\alpha(G)$ en una i.l.b., para obtener un resultado análogo al iii) de 2.6.2.

Sea f una i.l.b. en $S \subset \mathbb{R}^2$. Consideremos la familia $\{C_i\}$ $i \in M$ de las componentes isométricas de S , y la familia $\{f_i\}$ $i \in M$ de isometrías de \mathbb{R}^2 , tales que $\forall i \in M, f_i/C_i = f/C_i$. Consideremos el grupo H de isometrías de \mathbb{R}^2 generado por la familia $\{f_i\}$. Diremos que H es el grupo asociado a f .

Nos proponemos probar el siguiente resultado:

i) Si f es acíclica (no tiene elementos periódicos) y en el grupo H asociado a f no hay rotaciones de ángulo irracional, se verifica:

$$\alpha(G) = S.$$

Consideremos para probarlo el complementario $S - \alpha(G)$, y hagamos la hipótesis de que no es vacío. Teniendo en cuenta la definición del operador α podríamos encontrar un S -abierto $V \subset S - \alpha(G)$ en el que no hay puntos de la forma $f^{-k}(z)$ $k \geq 0, z \in G$, o lo que es equivalente, f^k/V es isometría. Podemos elegir V maximal tal que $\forall k \geq 0$ f^k/V es isometría y V es conexo. Pongamos $\forall k \in \mathbb{N}, V_k = f^k(V)$. Entonces $V_k \cap V_{k'}$ es vacío o coincide con V_k . Como los V_k tienen todos la misma medida que V , y es una medida positiva, si fueran todos disjuntos, S no sería acotado, de manera que para ciertos $k, k', k < k', V_k = V_{k'}$, y entonces, si $s = k' - k, f^s(V) = V$, y basta probar que eso significa que, o bien en V hay puntos periódicos para f , o bien en H hay una rotación irracional.

Tendremos también que $f^{2s}(V) = V$ donde f^{2s} es un movimiento directo del plano perteneciente al grupo H . Por tanto f^{2s} es un giro o una traslación. Esta última alternativa no es posible, puesto que para todo natural p , sería $f^{2sp}(V) = V$, y V no podría ser acotado. Si f^{2s} fuera un giro de ángulo racional, en V habría puntos periódicos, luego para que no los haya debe ser un giro de ángulo irracional, lo que también es imposible por hipótesis. ■

Obsérvese que si se permite la existencia en H de giros de ángulo irracional, prolongando el razonamiento, se vería que el giro no podría tener su centro en V , con lo que debe ser un giro con centro fuera de V , y V debe ser una corona circular. Así pues obtenemos:

ii) Si f es acíclica, los únicos conjuntos ω -isométricos de interior no vacío y conexos que puede haber, son coronas circulares. En particular, no puede haber intervalos ω -isométricos. Si existen coronas circulares ω -isométricas, los ω -conjuntos de sus puntos son círculos.

Por otra parte, si S es una corona circular y f una rotación de ángulo irracional de centro en el centro de la corona, f es una i.l.b. en S , $G = \emptyset$, y S es ω -isométrico, de forma que este caso es posible.

Obsérvese también que en el grupo H asociado a un intercambio de intervalos en \mathbb{R}^2 , no hay sino traslaciones, y en el asociado a una reordenación por intervalos, generado por traslaciones y simetrías axiales de ejes horizontales y verticales y giros de π , tampoco hay giros irracionales, por lo que a todos los tipos de i.l.b. descritos en 4.1.4, puede aplicarse el resultado i).

4.2.4. El problema de la transitividad y de la minimalidad

Hasta aquí hemos expuesto los resultados de las i.l.b. en R que pueden ser extendidos a \mathbb{R}^2 con facilidad. Sin embargo los resultados sobre minimalidad en intercambios de intervalos, y de manera más general de reordenaciones por intervalos no pueden extenderse, y presentan una resolución que parece complicada, que requerirá, bien la introducción de condiciones suplementarias, bien resultados salvo casos excepcionales con probabilidad nula de los mismos. Veamos un caso en el que no pueden extenderse los resultados de R a \mathbb{R}^2 .

Tomemos por ejemplo un intercambio de intervalos T en el intervalo unidad de R , con sólo dos componentes isométricas que tengan longitudes α y $1-\alpha$, racionalmente independientes. Es sabido (ver ref. [K1]), que en estas condiciones el intercambio de intervalos es minimal, e incluso ergódico. Formemos ahora en \mathbb{R}^2 el intercambio de

intervalos producto $f = T \times T$ en el intervalo unidad de \mathbb{R}^2 , S (ver iii de 4.1.4). Este intercambio de intervalos en \mathbb{R}^2 , es del tipo descrito en 4.1.3. (ver fig 4.4), y puede reducirse a una traslación en el toro, que en este caso, por ser según el vector $(1-\alpha, 1-\alpha)$, da lugar, para cada x de S , a una órbita densa en una geodésica

cerrada del toro paralela a la diagonal del intervalo unidad en R^2 (ver fig. 4.5), y por tanto f no es minimal y ni siquiera transitiva. Obsérvese que aquí se cumple también la condición análoga a la regularidad de las i.l.b. en R que se exigía para obtener minimalidad en aquel caso.

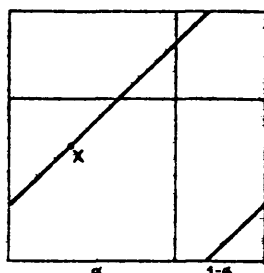


fig 4.5

4.2.5. Aplicaciones a billares poliédricos.

Para concluir el apartado, vamos a aplicar los resultados obtenidos al problema de los billares poliédricos racionales.

El teorema de recurrencia se traduce en la correspondiente propiedad recurrente de los billares poliédricos. Si S_θ es el espacio de fases asociado a una orientación θ , y $r : S_\theta \rightarrow S_\theta$ la correspondencia de reflexión, y μ la medida definida en S_θ descrita en 4.1.1, se obtiene :

i) μ -casi todos los puntos de S_θ son recurrentes.

En el caso del billar poliédrico, tampoco en el grupo H puede haber giros de ángulo irracional H , porque las transformaciones de H describen la actuación de las del grupo R generado por las simetrías respecto a las caras del poliedro, sobre las rectas $\{(P,B)\}$, donde $(P,B) \in S_\theta$ (ver 3.2.1). Una rotación $h \in H$, debe corresponder a una transformación $g \in R$ que produce una rotación de una recta r del tipo descrito. Entonces, la transformación g^2 , es un movimiento helicoidal que gira r , y puede descomponerse en producto de un giro axial j por una traslación t paralela al eje del giro. Si h fuera un giro de ángulo irracional, j también lo sería, y la actuación de g^2 sobre cualquier recta de orientación distinta a la del eje del giro, produciría una cantidad infinita de rectas de orientaciones distintas, lo que no es posible en un billar racional. Esto prueba:

ii) En un billar poliédrico, cualquier trayectoria de orientación θ tal que ningún (P, θ) tiene órbita periódica, pasa arbitrariamente cerca de alguna arista.

Ya que en esto se traduce el hecho de que $\alpha(G) = S_G$. Además, razonando como en el caso de billares poligonales racionales, vemos que si un elemento de trayectoria (P, θ) da lugar a una órbita periódica, hay un producto g de simetrías respecto a las caras del poliedro que mantiene invariante la recta $r = \{(P, \theta)\}$. La transformación g^2 será un movimiento helicoidal y fácilmente se ve que no puede ser la identidad. Ahora bien las únicas rectas invariantes en un movimiento helicoidal distinto de la identidad, tienen todas una misma orientación, que es la que corresponde al eje de la rotación y/o a la orientación de la traslación en que se descompone el movimiento helicoidal (salvo en el caso en que el movimiento helicoidal se reduzca a una rotación axial de ángulo π , en cuyo caso son invariantes todas las rectas perpendiculares al eje del giro, pero el cuadrado de dicho giro sería la identidad, lo que es imposible en nuestro caso). Por tanto, el número de orientaciones que dan lugar a órbitas periódicas, no puede exceder al de movimientos helicoidales contenidos en el grupo R , y como este grupo es finitamente generado y por consiguiente numerable, se obtiene:

iii) El conjunto de orientaciones θ tales que la órbita de algún (P, θ) es periódica, es numerable.

Este resultado da carácter genérico al anterior.

4.3. ALGUNOS RESULTADOS PARA ISOMETRIAS LOCALES PRODUCTO.

En este apartado se analizan casos particulares de reordenaciones por intervalos de \mathbb{R}^2 de los tipos expuestos en 4.1.4 en los que se pueden obtener resultados sobre la estructura de las órbitas.

4.3.1.

Tomemos dos reordenaciones por intervalos en \mathbb{R} , $T_1 : S_1 \rightarrow S_1$, y $T_2 : S_2 \rightarrow S_2$, sea $S = S_1 \times S_2$, y $T_1 \times T_2 = f : S \rightarrow S$. En estas condiciones:

i) Si $x = (x_1, x_2) \in S$ es tal que x_1 y x_2 son elementos periódicos respectivamente para T_1 y para T_2 , entonces x es periódico para f , y su período es el m.c.m. de los períodos de x_1 y x_2 .

ii) Si x es periódico para T_1 , S_2 es conexo y T_2 es regular y acíclica:

$$w_f(x) = (pr_1)^{-1}(w_{T_1}(x_1))$$

Donde pr_1 es la proyección sobre el eje horizontal. Como $w_{T_1}(x_1)$ es

por hipótesis un conjunto finito, $w_f(x)$ se compone de un número finito de segmentos copias verticales de S .

El primer resultado es evidente. El segundo se obtiene observando que si x_1 es periódico de periodo k , T_2^k es una reordenación por intervalos en S_2 regular y acíclica, y como S_2 consta de una sola componente conexa, minimal. Si $y_1 = T_1^P(x_1)$ e $y_2 = T_2^P(x_2)$ tomando $z = f^P(x) = (y_1, y_2)$, como $f^{mk}(z) = (y_1, T_2^{mk}(y_2))$, será $w_{f^k}(z) = \{y_1\} \times w_T^k(y_2) = \{y_1\} \times S_2$, y como $w_{f^k}(z) \subset w_f(x)$, tenemos que cada $\{T_1^P(x_1)\} \times S_2$, está incluido en $w_f(x)$. La inclusión contraria, es trivial. ■

4.3.2.

Consideramos ahora una reordenación por intervalos $T: S_1 \longrightarrow S_1$, donde S_1 es un intervalo real semicerrado, y suponemos T_1 regular y acíclica, siendo $\{J_i\}$, $1 \leq i \leq n$, la familia de componentes isométricas de T , y sea $\{T_i\}$, $1 \leq i \leq n$, una familia de intercambios de intervalos, $T_i: S_2 \longrightarrow S_2$, donde S_2 es el intervalo $[0,1]$ y T_i es para cada i un intercambio de dos intervalos semicerrados de longitudes $1-\alpha_i$ y α_i . Nos proponemos clasificar la reordenación por intervalos f en S obtenida por producto mixto de T por la familia $\{T_i\}$. (ver iv de 4.1.4 y fig 4.6). Equivalentemente, si añadimos su extremo superior a S_1 e identificamos ambos extremos, obteniendo un círculo S'_1 y hacemos lo propio con S_2 obteniendo un círculo S'_2 , f define una reordenación por intervalos en un toro $S' = S'_1 \times S'_2$ (ver 4.1.3), que consiste en cortar al toro en rodajas por planos normales al eje del toro e intercambiar estas rodajas de posición relativa, girando además cada una con un ángulo arbitrario. Por resultar mas cómodo este segundo modelo, cuya dinámica es equivalente a la del primero, haremos el estudio en el

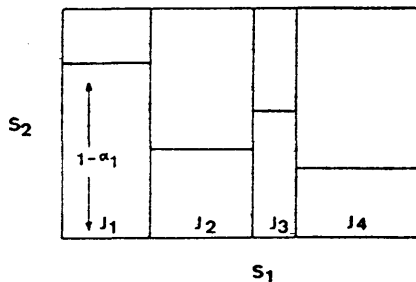


Fig 4.6

toro. La ventaja es la simplicidad de la expresión analítica de f . Si $x = (x_1, x_2)$ es un punto cualquiera de S' tal que $x_2 \in J_i$, entonces :

$$f(x_1, x_2) = (T(x_1), x_2 + \alpha_i)$$

ya que T_i es un giro de ángulo α_i en S'_2 .

Sea ahora $x = (x_1, x_2)$ e $y_1 \in S'_1$. Probaremos primero:

$$i) \omega_f(x) \cap \{y_1\} \times S'_2 \neq \emptyset$$

Como ambos conjuntos son compactos, si tuvieran intersección vacía su distancia sería positiva. Pero si $\epsilon > 0$, en la bola cerrada $B_\epsilon(y_1)$ debe haber infinitos puntos de la forma $T^k(x_1)$, debido a que, según las hipótesis hechas sobre T , este intercambio de intervalo es minimal. Ello significa que en el conjunto compacto $B_\epsilon(y_1) \times S'_2$ debe haber un valor de adherencia z para la sucesión de los $f^k(x)$ que pertenecen a dicho conjunto, y z es un punto de $\omega_f(x)$ tal que $d(z, \{y_1\} \times S'_2) \leq \epsilon$. Esta contradicción prueba el resultado. ■

Sea ahora $x' \in \omega_f(x) \cap \{x_1\} \times S'_2$, $x' = (x_1, x'_2)$. Entonces, si $\beta = x'_2 - x_2$ se tiene para todo $z = (x_1, z_2)$:

$$ii) (x_1, z_2 + \beta) \in \omega_f(z)$$

Para probarlo, tomemos una familia de bolas $\{V_i\}$, siendo V_i una bola centrada en x' , y de radio $1/i$, para todo $i \in \mathbb{N}$. En cada una de ellas podemos encontrar un elemento $t_i = f^{k_i}(x)$. Observemos además que en general, si para $1 \leq j \leq p$ es $f^{j-1}(x_1) \in J_m$, será :

$$f^p(x_1, z_2) = (T^p(x_1), z_2 + g), \quad g = \sum_{j=1}^p \alpha_{m_j}$$

por lo que si pr_2 representa la proyección sobre la segunda coordenada:

$$pr_2 \circ f^p(x_1, z_2) = z_2 + g$$

Siendo además g independiente de z_2 . Sea ahora g_i el valor de g correspondiente a $p = k_i$. Como la sucesión t_i converge a x' , la sucesión $pr_2 \circ f^{k_i}(z) = z_2 + g_i$ debe converger a $pr_2(x') = x'_2$, de donde g_i convergerá a $x'_2 - x_2 = \beta$. Esto significa que $pr_2 f^{k_i}(z) = z_2 + g_i$ convergerá a $z_2 + \beta$. ■

Supongamos ahora que x_1 pertenece al conjunto :

$$O = S_1 - \bigcup_i f^{-i}(X(T))$$

O lo que es equivalente, T^k es continua en x_1 para todo $k \in \mathbb{N}$. Sea

además $y = (x_1, y_2)$ y x' y β como en el caso anterior. Entonces:

$$iii) (x_1, y_2 + n\beta) \in \omega_f(y) \quad \forall n \in \mathbb{N}$$

Basta observar que f^k es continua en cada punto de $\{x_1\} \times S'_2$, ya que $pr_1 \circ f^k$ y $pr_2 \circ f^k$ son ambas continuas (la primera se escribe también $TK \circ pr_1$, y TK es siempre continua en x_1 , y la segunda se escribe $g \circ pr_2$, donde g es un giro, y es continuo en todo S'_2). Sabemos que $z = (x_1, y_2 + \beta) \in \omega_f(y)$ por ii), pero como f^k es continua en z , $\omega_f(z) \subset \omega_f(y)$, y aplicando ii) a z , obtenemos que $(x_1, y_2 + 2\beta) \in \omega_f(z) \subset \omega_f(y)$. Inductivamente, se consigue el resultado. ■

Supongamos ahora que $x = (x_1, x_2)$ cualquiera y sean $x_1 \in D$, e $y = (y_1, y_2) \in \omega_f(x)$. Sean x' y β como en ii). Entonces:

$$iv) (y_1, y_2 + n\beta) \in \omega_f(x), \quad n \in \mathbb{N}$$

Sea $\{t_i\}$ una sucesión tal que $t_i = f^{k_i}(x)$, $i \in \mathbb{N}$, y $t_i \rightarrow y$, sucesión que sabemos existe por ser $y \in \omega_f(x)$. Consideremos la sucesión $\{s_i\}$, donde $s_i = f^{k_i}(x')$, $i \in \mathbb{N}$. Sabemos que $t_i = (TK_i(x_1), x_2 + g_i)$ — ver ii) —, donde la primera coordenada tiende a y_1 , y la segunda a y_2 , para lo que g_i debe tender a $y_2 - x_2$. Esto significa que $s_i = (TK_i(x_1), x_2 + \beta + g_i)$, y por consiguiente, $s_i \rightarrow (y_1, y_2 + \beta)$. Este último punto pertenece a $\omega_f(x')$, y por la continuidad de f^k en $x' \quad \forall k$, pertenece también a $\omega_f(x)$. Reemplazando x por x' en el razonamiento, se obtendría el resultado para $(y_1, y_2 + 2\beta)$, y sucesivamente para cualquier n . ■

Probaremos ahora el resultado final.

v) f es minimal, o bien para todo $x \in S'$, $\omega_f(x)$ es un conjunto de medida nula.

Para demostrarlo, tomamos en primer lugar $x_1 \in D$. Consideramos cualquier $x = (x_1, x_2)$ y el conjunto $W = \omega_f(x) \cap \{x_1\} \times S'_2$, siendo $W \neq \emptyset$ por i). Sea $\beta = \inf\{x - y, y \in W - x\}$. Pueden darse dos alternativas excluyentes. La primera es que $m = 0$. La segunda es que m sea un número racional no nulo. No puede ser β un número irracional, porque según iii), todos los $(x_1, x_2 + k\beta)$ estarían en W . Ahora bien, por el teorema de Kronecker, esos puntos son densos en W , y m sería nulo. Veamos que sucede en los dos casos posibles:

a) Si $m = 0$, debe ser $W = \{x_1\} \times S'_2$, ya que podemos encontrar puntos $x' \in W$ tales que $\beta' = x - x'$ sea lo suficientemente pequeño para que los $(x_1, x_2 + k\beta')$ penetren en cualquier abierto de $\{x_1\} \times S'_2$, y dichos puntos pertenecen a W . Así W denso en $\{x_1\} \times S'_2$ y ambos conjuntos coinciden. Utilizando ii) se deduce además que esto sucede independientemente del valor de x_2 . No solo esto, sino que además, si $y_1 \in D$, $\omega_f(x) \cap \{y_1\} \times S'_2$ es denso en $\{y_1\} \times S'_2$, ya que en este conjunto hay algún $y = (y_1, y_2) \in \omega_f(x)$, y aplicando el resultado iv) se obtiene el resultado.

Probaremos ahora que para cualquier $z \in S'$, $\omega_f(z) = S'$. En $\{x_1\} \times$

S'_2 debe haber puntos de $\omega_f(z)$, según i). Sea x uno de ellos. Sabemos que $\omega_f(x) \cap [x_1] \times S'_2$ es denso en $[x_1] \times S'_2$ por la hipótesis hecha en el caso a) que nos ocupa. Si tomamos un abierto U cualquiera de S' , en U hay puntos $y = (y_1, y_2)$ donde $y_1 \in D$, por ser D denso en S'_1 , y según se ha señalado mas arriba, $\omega_f(x) \cap \{y_1\} \times S'_2$ es denso en $\{y_1\} \times S'_2$. Por tanto $\omega_f(x) \cap U \neq \emptyset$. Pero como $x_1 \in D$, también $\omega_f(x) \subset \omega_f(z)$, de donde $\omega_f(z) \cap U \neq \emptyset$, y queda probada la densidad en el caso a).

b) $m = \beta$, donde β es racional no nulo. Es claro, aplicando iii), que $\omega_f(x) \cap [x_1] \times S'_2$ consta de puntos aislados, dos consecutivos de los cuales se diferencian en β , y esto sucede con independencia del punto x elegido en $[x_1] \times S'_2$. Aplicando iv) se obtiene que pasa lo mismo con todos los $\omega_f(x) \cap \{y_1\} \times S'_2$, con $y_1 \in D$. Pero si z no pertenece a D , para un k finito, $x = f^k(z) \in D$, y según vi) de 2.6.1, $\omega_f(z) = \omega_f(x)$. Por tanto, para cualquier z de S' , el conjunto $\omega_f(z)$ tiene medida nula, por tener sus secciones verticales cardinal finito, excepto tal vez una cantidad numerable de ellas — las correspondientes a puntos de abscisa y_1 no perteneciente a D —. ■

Con ello concluye la clasificación de este tipo de transformaciones del toro.

4.3.2. Algunas definiciones.

Antes de proseguir, recordamos algunas definiciones clásicas en teoría ergódica, e introducimos el concepto de automorfismo semimezclador.

i) **Definición de aplicación ergódica.** Sea (X, \mathfrak{A}, μ) un espacio de probabilidad, y sea $T : X \rightarrow X$ una aplicación medible que preserve la medida (tal que para todo $A \in \mathfrak{A}$ es $\mu(T^{-1}(A)) = \mu(A)$). Se dice que T es ergódica si cada $A \in \mathfrak{A}$ invariante por T (tal que la $\mu(A \Delta T^{-1}(A)) = 0$.) verifica $\mu(A) = 0$ o $\mu(A) = 1$.

ii) **Definición de automorfismo mezclador.** Sea (X, \mathfrak{A}, μ) un espacio de probabilidad y $T : X \rightarrow X$ un automorfismo (es decir, que preserva la medida, y es invertible, en el sentido de que existe $S : X \rightarrow X$ medible tal que para casi todo x es $T \circ S(x) = x$ y $S \circ T(x) = x$). Se dice que T es mezclador si para todos los $A, B \in \mathfrak{A}$, se verifica $\mu(T^{-n}(A) \cap B) \rightarrow \mu(A) \cdot \mu(B)$ cuando $n \rightarrow \infty$.

Es fácil comprobar que si T es mezclador, es ergódico, ya que si A es invariante, para todo n es $\mu(T^{-n}(A) \cap A) = \mu(A)$, por lo que se tendrá $\mu(A) = (\mu(A))^2$, y debe ser $\mu(A) = 0$ o $\mu(A) = 1$.

Sin embargo, si S es el círculo y $T : S \rightarrow S$ es $T(x) = x + \alpha$ con α irracional, T es ergódico y no es mezclador.

Introduciremos aquí una propiedad intermedia entre estas dos:

iii) **Definición de automorfismo semimezclador.** Sea (X, \mathfrak{A}, μ) un espacio de probabilidad y T un automorfismo en \mathfrak{A} .

Decimos que T es semimezclador si $\forall A, B \in \mathcal{A}$ con $\mu(A), \mu(B) > 0$, existe n_0 tal que $n > n_0 \Rightarrow \mu(T^{-n}(A) \cap B) > 0$.

Nota : Existe ya (ver ref. [M1]) el concepto de automorfismo débilmente mezclador. La condición exigida es :

$$\lim_{n \rightarrow \infty} \frac{1}{n} \left| \sum_{i=0}^{n-1} \mu(T^{-i}(A) \cap B) - \mu(A) \cdot \mu(B) \right| = 0$$

La consecuencia de la condición de ergodicidad, es la proporcionalidad entre la porción de órbita de cada elemento que yace en un conjunto y la medida del mismo. La consecuencia de la condición de mezclador es la distribución uniforme de los elementos de un conjunto A en todo el espacio por las iteraciones de T . La condición de semimezclador significa que los conjuntos deben distribuirse por todo el espacio, aunque no necesariamente de forma uniforme. Es evidente que si T es mezclador, es semimezclador. También se ve fácilmente que si T es semimezclador es ergódico, ya que si A es un conjunto invariante con $\mu(A) \neq 1, 0$, los conjuntos A y $B = X - A$ no verifican la condición exigida en la definición iii). Sin embargo el ejemplo dado más arriba para probar que existen automorfismos ergódicos no mezcladores, prueba que existen automorfismos ergódicos no semimezcladores.

4.3.3.

Probaremos aquí una propiedad interesante de las reordenaciones por intervalos semimezcladoras. Supongamos que $K = \{k_i\}$, $i \in \mathbb{N}$, es una sucesión estrictamente creciente de números naturales. Sea f una reordenación de intervalos en S . Definimos para cualquier x de S :

$$\omega_{f,K}(x) = \bigcap_{n=0}^{\infty} \bigcup_{i=n}^{\infty} f^{k_i}(x)$$

Afirmamos entonces que si f es un automorfismo semimezclador en S , siendo aquí μ la medida Lebesgue normalizada, se verifica:

$$i) \mu \{ x \in S : \omega_{f,K}(x) \neq S \} = 0.$$

Llamemos A al conjunto encerrado entre las llaves. Si $x \in A$, podemos encontrar un $y \in S$ tal que y no está en $\omega_{f,K}(x)$, y por consiguiente un entorno de y , V_y , en el que no hay puntos de la forma $f^{k_i}(x)$. Si consideramos una base numerable $\{B_i\}$ de la topología usual en S , podemos encontrar cierto B_n de dicha base incluido en V_y . Esto prueba que si A_n es el conjunto de los puntos x de S tales que ningún $f^{k_i}(x)$ está en B_n , entonces $A = \bigcup A_n$ (la inclusión $\bigcup A_n \subset A$ es evidente). El conjunto A_n es medible, porque se puede escribir:

$$A_n = S - \bigcup f^{-k_i}(B_n)$$

Siendo $f^{-K_i}(B_n)$ medible por ser f una reordenación por intervalos. Esto prueba que A también es medible. Además $\mu(A_n) = 0$ para todo n , ya que en caso contrario entraríamos en contradicción con la hipótesis hecha de que f es un automorfismo semimezclador. En efecto si A_n tiene medida positiva, debe existir un n_0 tal que si $i > n_0$ sea $\mu(f^{-i}(B_n) \cap A_n) > 0$, lo que exige en particular que sea $f^{-i}(B_n) \cap A_n \neq \emptyset$. Sin embargo la sucesión K_i toma valores K_i arbitrariamente grandes (por ser estrictamente creciente), y sabemos que para cualquier i es $A_n \cap f^{-K_i}(B_n) = \emptyset$. Por tanto A_n tiene medida nula, y consiguientemente A también. ■

4.3.4.

Con el resultado del punto anterior se puede probar el siguiente resultado:

i) Si f_2 es una reordenación por intervalos semimezcladora en S_2 y f_1 es una reordenación por intervalos minimal en S_1 , entonces $f = f_1 \times f_2$ es casi minimal en $S = S_1 \times S_2$ (es decir, para casi todo $x \in S$ es $\omega_f(x) = S$).

Tomemos un $y = (y_1, y_2)$ cualquiera de S , y sea $x_1 \in S_1$. Como f_1 es minimal, existe una sucesión $K = \{K_i\}$ estrictamente creciente tal que $f_1^{K_i}(x_1) \rightarrow y_1$. Para casi todos los $x_2 \in S_2$ se verifica $\omega_{f,K}(x_2) = S_2$, y en particular, que y_2 es valor de adherencia de la sucesión $f_2^{K_i}(x_2)$, y por tanto, para casi todos los $(x_1, x_2) \in \{x_1\} \times S_2$, (y_1, y_2) es valor de adherencia de $\omega_{f,K}(x_1, x_2)$, e $y \in \omega_f(x_1, x_2)$. El conjunto de puntos x de S tales que y no pertenece a $\omega_f(x)$ tiene medida nula, por tener medida nula todas sus secciones verticales.

Consideremos el conjunto A de los $x \in S$ tales que $\omega_f(x) \neq S$. Para cada x de S hay un B_n perteneciente a una base numerable de abiertos $\{B_n\}$ tal que $B_n \cap \omega_f(x) = \emptyset$. (basta tomar un B_n incluido en el abierto $S - \omega_f(x)$). A se pone como unión de conjuntos $A_n = \{x \in S : B_n \cap \omega_f(x) = \emptyset\}$. Pero si seleccionamos en A_n un punto y_n , $A_n \subset C_n = \{x \in S : y_n \cap \omega_f(x) = \emptyset\}$, y sabemos por lo dicho arriba que $\mu(C_n) = 0$, de donde $\mu(A_n) = 0$, y $\mu(A) = 0$. ■

El resultado anterior puede ser utilizado en dos direcciones: Bien para probar la minimalidad de reordenaciones por intervalos en \mathbb{R}^2 obtenidas por producto cartesiano, si se consigue probar que los factores cumplen la hipótesis, o bien para probar que una reordenación por intervalos no tiene la propiedad semimezcladora. En particular se obtiene el siguiente resultado:

ii) Una condición suficiente para que una reordenación por intervalos f no sea semimezcladora (ni por consiguiente mezcladora) es que $f \times f$ no sea casi minimal.

Es consecuencia inmediata de la anterior. Esto hace particularmente interesante el estudio de las reordenaciones por intervalos $f \times f$, con vistas a la solución del problema de la determinación de si f

es mezcladora o no, problema para el que, hasta donde sabemos, no hay respuesta. Un resultado que se puede probar, es el siguiente:

iii) Si f es una reordenación por intervalos minimal en S_1 , y g es la reordenación por intervalos $f \circ f$ en $S = S_1 \times S_1$, el conjunto:

$$A = \{ x \in S : \omega_g(x) \neq S \}$$

Es denso en S .

Para probarlo, llamemos Γ_n al gráfico de la reordenación por intervalos $f^n : S_1 \rightarrow S_1$. Tomemos un punto $x \in \Gamma_n$. Tal punto tendrá coordenadas $(x_1, f^n(x_1))$, y su imagen por g tendrá coordenadas $(f(x_1), f^{n+1}(x_1))$, lo que significa que también es un punto de Γ_n . Como Γ_n es unión finita de segmentos, no puede ser denso en S , y $\omega_f(x)$ está incluido en la adherencia de G_n , queda probado que $\omega_f(x) \neq S$, y $\Gamma_n \subset A$. Solo resta probar que $\bigcup \Gamma_n$ es denso en S . Esto es cierto, porque si $x = (x_1, x_2) \in S$, por la minimalidad de f podemos encontrar un k tal que $f^k(x_1)$ esté arbitrariamente próximo a x_2 , y por tanto el punto $(x_1, f^k(x_1)) \in \Gamma_k$ está arbitrariamente próximo a x . ■

Nota : Señalemos además que si f es regular y $x = (x_1, x_2) \in \Gamma_n$, $\omega_g(x)$ coincide con la adherencia de Γ_n , porque f^n será una reordenación por intervalos, por lo que gráfico no puede tener puntos aislados. Un entorno U de un punto y de adherencia al gráfico, lo corta en cierto segmento s no reducido a un punto. Si s' es la proyección de s sobre el eje horizontal, por la minimalidad de f podemos encontrar $k > n$ tal que $f^k(x_1) \in s'$, el punto $(f^k(x_1), f^{k+n}(x_1)) \in s \subset U$ es un punto de Γ_n .

No obstante, $\mu(\bigcup \Gamma_n) = 0$, ya que cada Γ_n tiene medida nula.

En las fig. 4.7 y 4.8 se muestra el gráfico de una r.i. y el gráfico de su iterada 100.000., obtenidos por computador mediante esta idea: Con un programa que sirve para iterar $f \circ f$, se buscan respectivamente puntos de coordenadas $(x, f^k(x))$ con $k = 1$ y $k = 10^5$ y luego se itera la $f \circ f$ de ese punto hasta que cubre densamente el gráfico de f^k .

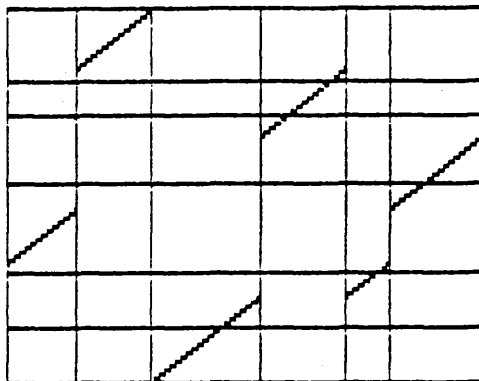


fig 4.7



fig 4.8

4.4. BILLARES DE POLIGONALES DE ANGULOS IRRACIONALES.

En este apartado se muestra como puede construirse un espacio de fases para el estudio de los billares poligonales de ángulos cualesquiera, de manera que la transformación del espacio de fases que describe el movimiento de las partículas en el billar tiene algunas características comunes con las isometrías locales en R^2 , en particular la conservación de la medida, lo que permite asegurar que este tipo de billares también es recurrente.

4.4.1. Espacio de fases.

Adoptamos para el billar poligonal las notaciones de 3.1.1. Si el billar no es racional, el número de diferentes orientaciones que puede adoptar una trayectoria no es finito, y no puede ser reducido el problema del billar a una reordenación por intervalos en R por lo que procederemos a construir un espacio de fases diferente.

Si ∂Q es el borde del billar, que para mayor simplicidad supondremos conexo, nuestro espacio de fases estará formado por el producto cartesiano $M = \partial Q \times (0, \pi)$. Cada punto (P, θ) de M puede ser interpretado como un elemento de trayectoria que parte de P y forma un ángulo θ con el lado de ∂Q al que pertenece P . Para que esté bien definido este ángulo, es necesario orientar previamente ∂Q , ya que ahora θ no mide el ángulo del elemento de trayectoria con una semirrecta fija, como en el caso de los billares racionales, sino con el lado a que pertenece P , que debe estar por tanto orientado. A tal fin bastaría establecer un sentido positivo para el recorrido de la poligonal ∂Q — el sentido antihorario, por ejemplo, — con lo que tendremos definido un sentido en cada lado. Además, los ángulos que forman los elementos de trayectoria que parten de puntos de ∂Q hacia el interior de la mesa con los lados del billar orientados de esta manera, son ángulos mayores que 0 y menores que π , por lo que todos los elementos de trayectoria están incluidos en S . Finalmente si tomamos cada uno de los lados como un segmento cerrado por el extremo que corresponde al origen (origen respecto al sentido elegido en el lado), y abierto en el extremo contrario, cada vértice pertenecerá a un único lado, con lo que el ángulo de las trayectorias partiendo de los vértices estará únicamente definido.

La aplicación de reflexión $r: M \rightarrow M$ se define como en 3.1.1, y hace corresponder a cada elemento de trayectoria su reflejado en el lado en que incide. Introduciremos ahora un sistema de coordenadas en S , para luego estudiar la expresión de r respecto a este sistema de coordenadas. Sea $\partial Q = \{l_i\}$, $1 \leq i \leq m$, donde los l_i son los lados de la mesa ordenados según el sentido de recorrido. Comenzamos definiendo una coordenada local x_i en cada lado l_i tal que $x_i(P)$ mide la distancia del punto $P \in l_i$ a su origen O_i . Establezcamos coordenadas locales en $M_i = l_i \times (0, \pi)$, haciendo corresponder a cada $(P, \theta) \in S$ el punto $(x_i(P), \theta)$ del plano. Finalmente si a_i representa la longitud de l_i y definimos:

$$c_i(\theta) = \sum_{j=1}^{i-1} a_j \cdot \sin \theta$$

La aplicación $\varphi : M \longrightarrow \mathbb{R}^2$ definida por:

$$\varphi(P, \theta) = (c_i(\theta) + x_i(P), \theta) \quad \forall (P, \theta) \in I_i, 1 \leq i \leq m$$

Establece una biyección entre M y cierta región del plano S . Cada sección $\partial Q \times \{\theta\}$ se aplica biyectivamente sobre una sección horizontal de S de la forma $[0, \lambda_\theta]$ donde:

$$\lambda_\theta = \sum_{i=1}^m a_i \cdot \sin \theta$$

Esta aplicación es continua en todos los puntos de M salvo en los puntos de la forma (O_i, θ) . Téngase en cuenta que $\partial Q \times \{\theta\}$ es homeomorfo a una circunferencia, $[0, \lambda_\theta]$ es un intervalo semicerrado, y (O_i, θ) el punto por donde φ "rompe" la circunferencia para aplicarla sobre el intervalo.

Pero lo interesante de esta construcción es que la biyección φ induce sobre S una transformación que conserva la superficie. Tal transformación viene definida por $f = \varphi \circ \tau \circ \varphi^{-1}$. Tomemos como en la figura 4.7 un punto (P, θ) en I_1 , y sea $\tau(P, \theta) = (Q, \theta') \in I_2$. Si α_{12} expresa el ángulo formado por estos dos lados, se tiene $\alpha_{12} = \theta + \theta'$, de donde:

$$\theta' = \alpha_{12} - \theta$$

Por otra parte, al punto (P, θ) le corresponden coordenadas

$(x_1(P) \sin \theta, \theta)$, y al punto $\tau((P, \theta)) = (Q, \theta')$ le corresponden coordenadas $(x_2(Q) \sin \theta', \theta')$. De la figura se deduce:

$$(a_1 - x_1(P)) / \sin \theta' = x_2(Q) / \sin \theta$$

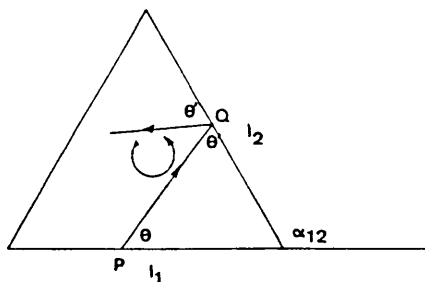


fig. 4.7

lo que significa que $x_2(Q)\text{sen}\theta' = a_1\text{sen}\theta - x_1(P)\text{sen}\theta$. Así, un punto de coordenadas (x, θ) de M_1 se transforma por f en un punto de coordenadas $(a_1\text{sen}\theta - x, \alpha_{12} - \theta)$, donde a_1 y α_{12} son independientes del punto elegido en l_1 . En general un punto (x, θ) correspondiente al lado l_i que se aplica por f sobre un punto correspondiente al lado l_j , se transforma por f en un punto de coordenadas $(a_i\text{sen}\theta - x, \alpha_{ij} - \theta)$, siendo a_i y α_{ij} constantes para todos los puntos en las citadas condiciones. Esta transformación es continua en todos los puntos, salvo en las preimágenes por f de puntos de S correspondientes a vértices. Fácilmente se ve que dichas preimágenes corresponden a los conjuntos imágenes de una colección finita de curvas continuas, cuya unión llamaremos G . Fuera de ellas, f es diferenciable con continuidad. La matriz jacobiana de la transformación f es:

$$\begin{bmatrix} -1 & 0 \\ a_i \cos \theta & -1 \end{bmatrix}$$

Lo que demuestra que f preserva el área, ya que el determinante de esta matriz vale siempre la unidad, y el conjunto de puntos de discontinuidad G tiene medida nula. Si consideramos S como un espacio de probabilidad con la medida de Lebesgue normalizada, f es un automorfismo en S que conserva la medida, y por tanto, podemos afirmar que, verificándose las hipótesis del teorema de Poincaré, casi todos los puntos de S son recurrentes, es decir, verifican que $x \in \omega_f(x)$.

Para finalizar el capítulo, señalamos la evidente analogía entre los requisitos cumplidos por $f: S \rightarrow S$ y la definición de isometría local biyectiva en R^2 , lo que según nuestro punto de vista reafirma el interés del futuro desarrollo de las técnicas que introducimos en los capítulos 2, 3 y 4 de esta memoria.

CAPITULO V

ALGORITMOS FRACTALES

5.0. INTRODUCCION.

En este capítulo se estudian ciertos algoritmos mediante los que se pueden generar conjuntos de estructura fractal.

El primer apartado del capítulo está destinado a exponer brevemente las nociones y resultados de teoría geométrica de la medida necesarios en el resto del capítulo.

El segundo apartado expone un procedimiento para la obtención de conjuntos fractales (en el sentido amplio utilizado en 5.1.4) mediante series de vectores de \mathbb{R}^n . Se asocia a cada serie absolutamente convergente un conjunto compacto, y se determinan las condiciones que debe cumplir la serie para que en su conjunto asociado quepa esperar propiedades fractales. Se llaman series fractales a las que cumplen estas condiciones, obteniéndose además una acotación superior para la dimensión de Hausdorff de sus conjuntos asociados.

El tercer apartado se dedica a desarrollar un estudio particular para las series geométricas. En el caso real, se prueba que la condición necesaria y suficiente para que el fractal asociado sea un conjunto autosemejante es que el valor absoluto de la razón sea menor que $1/2$, obteniéndose por tanto la dimensión de dichos fractales. En el caso complejo se prueba la suficiencia de la condición. Se analizan también conjuntos fractales caracterizados por las expresiones decimales de sus elementos, estableciéndose una interesante conexión entre dimensión de Hausdorff y el número de sumas diferentes que pueden ser obtenidas por una colección de naturales.

El cuarto apartado muestra como cada función analítica real o compleja, tiene asociada toda una familia de fractales (uno por cada punto de ciertos entornos), que además tendrán en general diferentes dimensiones y aspectos muy variados, como se muestra en el anexo de gráficos. Se obtienen las dimensiones de dichos fractales para todas las funciones de la forma $f(x) = (1-x)^{-p}$, para $p \in \mathbb{N}$.

Finalmente en el último apartado se muestran otros algoritmos, que, con espíritu parecido a las series fractales, permiten también la construcción de conjuntos fractales, como los productos infinitos, fracciones continuas y otros algoritmos que permiten la codificación y obtención gráfica de fractales conocidos, como el conjunto de Besicovitch, ciertos autosemejantes, etc.

En este capítulo el símbolo $|x|$ se utiliza indistintamente para diámetro si x es un conjunto y para norma si es un vector. Esperamos que el contexto evite toda posibilidad de confusión.

5.1. DEFINICIONES PREVIAS.

5.1.1. Medidas de Hausdorff

En 1919, Hausdorff introdujo un sistema de medidas mas fino que la medida de Lebesgue, en el sentido de que permite diferenciar entre sí conjuntos de muy diferente naturaleza y propiedades geométricas que no pueden ser separados por la medida de Lebesgue.

Partimos de un espacio métrico E y de un subconjunto F de E . Fijamos números reales $s \geq 0$ y $\delta > 0$, y consideramos la clase C_δ de los δ -recubrimientos de F , es decir de los recubrimientos de F por conjuntos de diámetros menores que δ . Si $|C_i|$ expresa el diámetro del conjunto C_i definimos:

$$H^s_\delta(F) = \inf \sum |C_i|^s$$

Donde el ínfimo se toma sobre todos los recubrimientos de C_δ . Como si es $\delta' < \delta$, resulta $C_{\delta'} \subset C_\delta$, al disminuir δ el ínfimo aumenta, por lo que :

$$\lim_{\delta \rightarrow 0} H^s_\delta(F) = \sup_{\delta > 0} H^s_\delta(F)$$

Esto permite definir:

$$H^s(F) = \lim_{\delta \rightarrow 0} H^s_\delta(F)$$

Y la función de conjuntos así definida es la medida de Hausdorff en dimensión s , o H^s -medida. Se trata de una medida exterior métrica, en cuya σ -álgebra de conjuntos medibles están incluidos todos los conjuntos de Borel de E . La restricción de H^s a dicha σ -álgebra es una medida regular.

5.1.2. Dimensión de Hausdorff

En general, en el sentido que explicamos a continuación, $H^s(F)$ será cero o infinito si no seleccionamos adecuadamente s . En efecto, existe para cada $F \subset E$ un único número real que se llama dimensión Hausdorff de F y se escribe $\dim(F)$ tal que si $s < \dim(F)$, $H^s(F) = \infty$ y si $s > \dim(F)$, $H^s(F) = 0$. Nada se puede afirmar de la medida de Hausdorff de F en su propia dimensión, que puede ser cero, infinito, o un número positivo. En este último caso, si $s = \dim(F)$ se dice que F es un s -conjunto.

Si $E = \mathbb{R}^n$ se verifica:

$$i) \dim(F) \leq n$$

$$ii) G \subset F \Rightarrow \dim(G) \leq \dim(F)$$

Es fácil comprobar que si $E = \mathbb{R}^n$, la medida Hausdorff en dimensión n es un múltiplo de la medida de Lebesgue:

$$H^n(F) = c_n \cdot L_n(F)$$

Donde L_n es la medida Lebesgue n -dimensional y c_n es una constante que no depende mas que de n .

No obstante, es frecuente el caso en que F tiene medida Lebesgue nula y sin embargo, la medida Hausdorff de F en su propia dimensión Hausdorff es positiva. Tal es por ejemplo el caso del conjunto C de Cantor, que tiene medida Lebesgue lineal nula y cuya dimensión de Hausdorff es $s = \log 2 / \log 3$, siendo $H^s(C) = 1$.

5.1.3. Medidas de Hausdorff generalizadas

Un sistema más fino de medidas puede ser construido según el procedimiento descrito por Rogers (ver ref. [R1]). Este procedimiento permite asociar a cada función g de una clase muy amplia una medida H_g . Consideramos para ello funciones g de variable real no negativa, con $g(t)$ positivo si t es positivo, monótonas crecientes y continuas a la derecha en todos los puntos. Manteniendo las notaciones de 5.1.1 definimos:

$$H_g(F) = \liminf_{\delta \rightarrow 0} \sum g(|C_i|)$$

Donde el ínfimo se toma sobre todos los δ -recubrimientos de F . La función de conjuntos así definida es una medida exterior, con todas las propiedades de las medidas usuales de Hausdorff, que no son sino un caso particular de este proceso cuando $g(t) = t^s$.

5.1.4. Conjuntos fractales y autosemejantes.

La definición más común de conjunto fractal es la propuesta por Mandelbrot : Un conjunto fractal es aquel cuya dimensión de Hausdorff es estrictamente superior a su dimensión topológica, siendo también un hecho admitido que esta definición no es enteramente satisfactoria por quedar al margen de ella ciertos conjuntos cuya estructura es comunmente admitida como fractal. Mandelbrot proporciona (ver ref. [M3]) la siguiente definición, matemáticamente vaga, pero que da cierta idea intuitiva de la estructura fractal de un conjunto: " Fractal. adj. Sentido intuitivo. Que tiene una forma bien sea sumamente irregular, bien sumamente interrumpida o fragmentada y sigue siendo así a cualquier escala que se produzca el examen. Que contiene elementos distintivos cuyas escalas son muy variadas y cubren una gama muy amplia."

Uno de los ejemplos más importantes de conjuntos fractales son los denominados autosemejantes, entre los que se encuentran varios de los más conocidos: Conjunto de Cantor, curva de Koch (copo de nieve), triángulo de Sierpinsky etc.

Para definir conjunto autosemejante, consideramos una colección finita de semejanzas en R^n , $\{\psi_i\}$, $1 \leq i \leq m$, tales que la razón r_i de ψ_i sea $r_i < 1$, $1 \leq i \leq m$. A partir de esta familia se define una transformación de conjuntos en R^n por la fórmula :

$$\Psi(F) = \bigcup_{i=1}^m \psi_i(F)$$

En estas condiciones, un conjunto $F \subset \mathbb{R}^n$ es autosemejante si verifica:

$$a) \Psi(F) = F$$

$$b) \exists s: H^s(F) > 0 \text{ y } H^s(\psi_i(F) \cap \psi_j(F)) = 0, 1 \leq i \neq j \leq m$$

La primera condición siempre resulta verificada por cierto subconjunto compacto F de \mathbb{R}^n , que sólo depende de la transformación Ψ definida por la familia de semejanzas. De hecho, se demuestra que para una cierta distancia definida entre los subconjuntos compactos de \mathbb{R}^n , llamada la métrica de Hausdorff, dichos subconjuntos compactos forman un espacio métrico, y Ψ es una aplicación contractiva en este espacio, por lo que se verifica el siguiente teorema:

i) existe un único punto fijo F para Ψ . Además, si G es un conjunto compacto arbitrario (no vacío) de \mathbb{R}^n , las iteraciones $\Psi^n(G)$ convergen a F en la métrica de Hausdorff.

Los conjuntos autosemejantes se caracterizan porque la totalidad del conjunto puede reconstruirse a partir de cualquier fragmento (no trivial) del mismo por una simple semejanza. La segunda condición es en general de difícil verificación. Expresa que las copias de F por las semejanzas de la familia no solapan excesivamente entre sí. Esta condición puede ser sustituida por la útil condición de abierto, que además proporciona con exactitud la dimensión de Hausdorff del conjunto autosemejante. En efecto, se verifica el siguiente teorema:

ii) Si existe un conjunto abierto y acotado V tal que $\Psi(V) = \bigcup_{i=1}^m \psi_i(V)$, siendo esta unión disjunta, entonces el único conjunto F tal que $\Psi(F) = F$ es autosemejante, y su dimensión de Hausdorff es el único número real s que verifica:

$$\sum_{i=1}^m (r_i)^s = 1$$

La hipótesis de este teorema, es la llamada *condición de abierto*.

5.2. SERIES FRACTALES

5.2.1. Conjunto asociado a una serie

Sea $\sum_{i=0}^{\infty} a_i$ una serie absolutamente convergente, donde $a_i \in \mathbb{R}_n$.
 Sea también el conjunto \mathbb{N}^* de los naturales mas el cero, y sea $P(\mathbb{N}^*)$ el conjunto de las partes de \mathbb{N}^* . Construyamos una aplicación $\nu_a : P(\mathbb{N}^*) \rightarrow \mathbb{R}^n$ definida de la siguiente manera:

$$\nu_a(\emptyset) = 0$$

$$\nu_a(I) = \sum_{i \in I} a_i, \text{ si } I \in P(\mathbb{N}^*), I \neq \emptyset$$

De esta forma ν_a es una función numerablemente aditiva en $P(\mathbb{N}^*)$, y de hecho si $a_i \in \mathbb{R}$ se trata de una medida real, y si $a_i \in \mathbb{C}$, de una medida compleja.

Consideremos el conjunto imagen por ν_a de $P(\mathbb{N}^*)$:

$$C_a = \nu_a(P(\mathbb{N}^*)) = \bigcup_{I \subset \mathbb{N}^*} \sum_{i \in I} a_i$$

C_a , según esta definición, es el conjunto de todas las posibles sumas de términos de la serie $\sum a_i$.

Probaremos que bajo ciertas condiciones exigidas a la serie $\sum a_i$, el conjunto C_a tiene propiedades fractales. A las series que cumplan estos requisitos las llamaremos series fractales.

5.2.2. Propiedad de compacidad.

El conjunto C_a está acotado, ya que se encuentra incluido en una bola de centro en el origen y radio $\sum |a_i|$.

Demostraremos a continuación que C_a es un conjunto cerrado. Sea $\{I_p\}$ una sucesión de partes de \mathbb{N}^* , de forma que $\nu_a(I_p)$ sea una sucesión convergente tal que $\nu_a(I_p) \rightarrow x \in \mathbb{R}^n$. Debemos demostrar que $x \in C_a$.

Vamos a construir para ello un conjunto J de índices tal que $\nu_a(J) = x$, definiendo inductivamente los elementos j_0, j_1, \dots, j_n de J .

Supongamos que $\forall i \in \mathbb{N}^*$, i pertenece tan sólo a un número finito de miembros de $\{I_p\}$. Fijemos un i cualquiera y descartemos los finitos miembros de $\{I_p\}$ a los que pertenece i o alguno de los números naturales anteriores al i . La sucesión restante, $\{H_i, p\}$, verifica que $\forall p \mid \nu_a(H_i, p) \rightarrow R_i$ donde:

$$R_i = \sum_{n \geq i} |a_n|$$

Además es claro que todavía se verifica: $\mu_a(H_{i,p}) \rightarrow x$. Tomando i suficientemente avanzado, podemos conseguir, por la convergencia absoluta de la serie $\sum a_i$, que R_i sea tan pequeño como queramos. De ello y de la acotación arriba señalada, se deduce que $x = 0$, y el conjunto de índices buscado es $J = \emptyset$.

Supongamos ahora que no es vacío el conjunto de índices i de N^* para los que i está en infinitos I_p . Sea j_0 el mínimo de los índices que verifican dicha condición, de tal manera que ninguno de los naturales anteriores a j_0 están en infinitos I_p . Ello significa que si suponemos los I_p ordenados según el orden natural en N^* , infinitos de los I_p deben comenzar en j_0 . Descartando los que no lo hagan, obtenemos una subsucesión de conjuntos que comienzan por j_0 y que llamaremos también $\{I_p\}$. A partir de ella construimos una nueva sucesión, la $\{I_{1,p}\}$, tal que $\forall p$ $I_{1,p}$ se obtiene suprimiendo de I_p j_0 . Tendremos:

$$a) \lim_{p \rightarrow \infty} \mu_a(I_{1,p}) = x - a_{j_0}$$

$$b) |\mu_a(I_{1,p})| \leq R_{j_0}$$

Procedemos a continuación por inducción. Supongamos definidos j_0, j_1, \dots, j_{n-1} e $\{I_{n,p}\}$ verificando:

$$a) \lim_{n \rightarrow \infty} \mu_a(I_{n,p}) = x - a_{j_0} - a_{j_1} - a_{j_2} - \dots - a_{j_{n-1}}$$

$$b) |\mu_a(I_{n,p})| \leq R_{j_{n-1}}$$

Si no hay ninguno de los a_j que pertenezca a infinitos de los $I_{n,p}$, razonando como en el primer paso de la inducción, se concluye que $x - a_{j_0} - a_{j_1} - \dots - a_{j_{n-1}} = 0$, por lo que para $J = \{j_0, j_1, \dots, j_{n-1}\}$ se consigue $\mu_a(J) = x$.

Supongamos que algún i pertenece a infinitos $I_{n,p}$. Sea j_n el mínimo con dicha propiedad. Descartando los $I_{n,p}$ que contienen elementos anteriores a j_n y quitando a los que restan j_n , se obtiene la sucesión $I_{n+1,p}$. Es claro que $\lim_{n \rightarrow \infty} \mu_a(I_{n+1,p}) = x - a_{j_0} - a_{j_1} - \dots - a_{j_n}$, y que $|\mu_a(I_{n+1,p})| \leq R_{j_n}$.

Queda así definido por inducción el conjunto de índices $J = \bigcup_{i=0}^{\infty} j_i$ para el cual probaremos que $\mu_a(J) = x$.

Dado $\epsilon > 0$, tomemos n suficientemente grande para que $R_j < \epsilon/2$.
 Para p suficientemente grande, si $y_n = x - a_{j_0} - a_{j_1} \dots - a_{j_n}$, se
 verifica $|y_n - \nu_a(l_{n+1}, p)| < \epsilon/2$. Poniendo:

$$|y_n| = |y_n - \nu_a(l_{n+1}, p) + \nu_a(l_{n+1}, p)| \leq \\
|y_n - \nu_a(l_{n+1}, p)| + |\nu_a(l_{n+1}, p)| < \epsilon/2 + \epsilon/2$$

Se prueba que $\lim_{n \rightarrow \infty} y_n = x - \lim_{n \rightarrow \infty} \sum_{i=0}^n a_{j_i} = x - \mu_a(J) = 0$. ■

5.2.3. Propiedad de bipartición

Sea para cualquier $j \in \mathbb{N}$ C_a^j el conjunto de todas las sumas de
 términos de la serie $\sum a_i$ formados con los términos del a_j en
 adelante. Se verifican las relaciones:

$$i) C_a = C_a^0 \\
ii) C_a = C_a^j \cup \{a_j + C_a^{j+1}\}, j \in \mathbb{N}^*$$

Sustituyendo en el segundo miembro de la igualdad i) la ii)
 particularizada para $j = 0$, se obtiene:

$$C_a = C_a^1 \cup \{a_0 + C_a^1\}$$

Sustituyendo en el segundo miembro de esta igualdad los C_a^1 por la
 ii) particularizada en $j = 1$, se obtiene:

$$C_a = C_a^2 \cup \{a_0 + C_a^2\} \cup \{a_1 + C_a^2\} \cup \{a_0 + a_1 + C_a^2\}$$

Reiterando el proceso se obtiene:

$$iii) C_a = \bigcup_{I \in P_j} \{ \sum_{i \in I} a_i + C_a^{j+1} \}, j \in \mathbb{N}^*$$

Donde P_j representa las partes del conjunto $\{0, 1, \dots, j\}$.
 Teniendo en cuenta que cada uno de los conjuntos encerrados en las
 llaves del segundo miembro de iii) es una copia por traslación del
 conjunto C_a^{j+1} , resulta el conjunto C_a descompuesto en unión de
 copias del C_a^{j+1} .

5.2.4. Definición de serie fractal

$\sum a_i$ es una serie normalmente convergente en R^n , donde , para evitar casos degenerados, suponemos que R^n coincide con el subespacio vectorial generado por los a_i , $i \in N^*$. Decimos que $\sum a_i$ es una serie fractal si:

i) C_a tiene potencia del continuo

ii) $L_n(C_a) = 0$.

De estas propiedades, y de la propiedad de bipartición, cabe esperar un comportamiento fractal para el conjunto C_a generado por la serie $\sum a_i$.

5.2.5. Condición necesaria para que una serie sea fractal.

Es evidente que para que $\sum a_i$ sea una serie fractal debe verificarse:

i) Para infinitos valores de i , es $a_i \neq 0$

Ya que en caso contrario, C_a sería un conjunto finito, y no se cumpliría la condición i) de serie fractal.

De esta condición necesaria, se desprenden dos consecuencias:

ii) Si se cumple i), C_a es un conjunto perfecto.

Podemos suponer para probarlo que la serie $\sum a_i$ consta de infinitos términos no nulos, suprimiendo si hace falta los términos nulos que inicialmente hubiera en la serie, ya que ello no modifica el conjunto C_a . Si $x \in C_a$ y es suma de infinitos términos de la serie, entonces x se obtiene como límite de sumas finitas de términos de la serie, y por tanto es de acumulación de C_a . Si x es suma de términos de la serie, entonces es límite de los $y_n \in C_a$, siendo:

$$y_n = x + \sum_{i>n} a_i$$

Y esto prueba que todos los puntos de C_a son de acumulación de C_a , y por tanto que C_a es perfecto, pues sabemos por 5.2.2 que C_a es cerrado. ■

iii) Si se verifica la condición i), C_a tiene potencia del continuo.

Podemos suponer para demostrarlo que $\sum a_i$ tiene todos sus términos no nulos, y que además están ordenados en orden decreciente de normas, ya que la convergencia absoluta de la serie garantiza que el conjunto C_a no varía por cambios de orden en los términos de la serie. Si conseguimos extraer de la serie $\sum a_i$ una serie $\sum b_i$ tal que el conjunto C_b tenga potencia del continuo, entonces se habrá probado que también la tiene C_a , puesto que $C_b \subset C_a$. Construyamos

$\sum b_i$ de forma inductiva, eligiendo $b_0 = a_0$, y si $b_n = a_k$, tomamos para b_{n+1} un a_i tal que $0 < R_i < |b_n|$. Tal elección es siempre posible, ya que $|b_n| > 0$ y $R_i \rightarrow 0$. La serie $\sum b_i$ verifica:

$$|b_i| > R'_i = \sum_{n>i} |b_n|$$

En efecto, si $b_{i+1} = a_r$, se tiene por construcción:

$$|b_i| > R_r = \sum_{n>r} |a_n| \geq \sum_{n>i} |b_n| = R'_i$$

Tomemos ahora dos conjuntos de índices I, J , tales que $I \neq J$. Probaremos que $\mu_b(I) \neq \mu_b(J)$ con lo que, teniendo $P(N^*)$ potencia del continuo, también la tendrá C_b .

Si $I \neq J$, entonces $I \Delta J = (I - J) \cup (J - I) \neq \emptyset$. Sea m el menor elemento de $I \Delta J$, y supongamos que $m \in I$. Tendremos:

$$\mu_a(I) - \mu_b(J) = \sum_{i \in I} b_i - \sum_{i \in J} b_i = b_m - \sum_{i \in J} \delta_i b_i$$

Donde $\delta_i = 0$ si $i \in I \cap J$, $\delta_i = -1$ si $i \in I$, y $\delta_i = 1$ si $i \in J$. La elección de m hace que los términos anteriores a b_m se anulen. Se tiene:

$$\begin{aligned} |\mu_a(I) - \mu_b(J)| &= |b_m - \sum_{i>m} \delta_i b_i| \geq |b_m| - |\sum_{i>m} \delta_i b_i| \\ &\geq |b_m| - \sum_{i>m} |\delta_i| |b_i| \geq |b_m| - \sum_{i>m} |b_i| \\ &= |b_m| - R'_m > 0. \end{aligned}$$

Esto concluye la demostración. ■

5.2.5. Condiciones suficientes para que una serie sea fractal

Para que la serie $\sum a_i$ sea fractal, es suficiente que:

i) Sea $a_i \neq 0$ para infinitos valores de i

ii) $2^i (R_i)^n \rightarrow 0$

La condición i) garantiza, según vimos en el punto anterior, que C_a tiene potencia del continuo. Basta comprobar que la condición ii) garantiza que $L_n(C_a) = 0$.

Según la propiedad de bipartición, concretamente según iii) de 5.2.3, podemos cubrir C_a por conjuntos de la forma $\sum_{i \in I} a_i + C_{a_{j+1}}$,

donde $I \subset P_j$. Todos estos conjuntos se obtienen por traslación del $C_{a_{j+1}}$. Todos tienen la misma medida Lebesgue, y como $\text{card}(P_j) = 2^{j+1}$, resulta:

$$L_n(C_a) \leq 2^{j+1} L_n(C_{a_{j+1}})$$

Ahora bien, el conjunto $C_{a_{j+1}}$ está incluido en una bola de centro en 0 y radio R_j , y por tanto $L_n(C_{a_{j+1}}) \leq c_n(R_j)^n$, donde c_n es una constante que sólo depende de n . por tanto:

$$L_n(C_a) \leq 2^{j+1} c_n(R_j)^n$$

Y si $2^j(R_j)^n \rightarrow 0$, resulta que $L_n(C_a) = 0$. ■

La condición ii) puede ser sustituida por la condición:

$$iii) 2^i R_i \rightarrow 0$$

que es independiente de n , ya que tomando i suficientemente grande, y teniendo presente la convergencia absoluta de la serie, $R^i > (R_i)^n$.

La condición ii), y desde luego la iii) no son en general necesarias. En efecto, veremos como en caso de repetición de los términos de la serie, estas condiciones pueden ser sustituidas por una condición más débil.

5.2.6. Función de repetición

Introducimos ahora una sucesión p_i asociada a las repeticiones de los términos de la serie $\sum a_i$. Suponemos que todos los términos de la serie son no nulos, y que están ordenados decrecientemente según sus normas. Definimos una sucesión auxiliar r_i de la siguiente manera:

$$a) r_0 = 2$$

$$b) \text{ si } i \neq 0 \text{ y } a_{i-k-1} \neq a_{i-k} = a_{i-k+1} = \dots = a_i, r_i = (k+2)/(k+1)$$

El efecto de esta definición, es que si un término a_i está precedido de un término distinto de él, $r_i = 2$, mientras que si está precedido de k términos consecutivos iguales, entonces:

$$r_{i-k} \cdot r_{i-k+1} \dots r_i = 2 \cdot 3/2 \cdot 4/3 \dots (k+2)/(k+1) = k+2$$

Se define ahora la sucesión p_i :

$$p_i = \prod_{n=0}^i r_n$$

Es claro que siempre se verifica que $p_i \leq 2^i$. La igualdad se produce en el caso de que la serie $\sum a_i$ no tenga términos repetidos. Afirmamos que las condiciones suficientes obtenidas en

el punto anterior, se pueden sustituir por las:

$$i) a_i \neq 0 \text{ para infinitos } i \in \mathbb{N}^*$$

$$ii) \rho_i(R_i)^n \rightarrow 0$$

Siendo esta última más débil que la ii) del punto anterior por ser $\rho_i \leq 2^i$.

Para comprobarlo, basta ver que en el recubrimiento de C_a por copias trasladadas de C_a descrito en iii) de 5.2.3, el número

de estas copias es menor o igual a r_j . Esto se debe a que las traslaciones de los C_a , que cuando todos los términos de la serie que preceden al a_j son diferentes, pueden ser como máximo $2j+1$, se repiten cuando hay repeticiones en los términos de la serie. Con un bloque de $k+1$ elementos iguales a uno dado, digamos al a_n , sólo pueden ser obtenidas $k+2$ traslaciones diferentes: la traslación nula, la a_j , la $2a_j$, etc, hasta la $(k+1)a_j$. El número máximo de traslaciones diferentes que pueden obtenerse con todos los términos de la serie anteriores a uno dado lo podemos computar agrupándolos en bloques de elementos repetidos, y multiplicando entre sí el número de las traslaciones que pueden obtenerse con los términos de cada bloque. Este cómputo coincide con la definición dada para la construcción de la función ρ_i de repetición. (El número exacto de traslaciones diferentes, puede ser todavía menor que ρ_i , ya que combinaciones de términos diferentes pueden tener sumas iguales, pero esta nueva economización parece difícil de llevar a cabo sin conocer la serie concreta de que se trate).

5.2.7. Acotación superior de la dimensión Hausdorff

Las condiciones suficientes para serie fractal obtenidas en el punto anterior, permiten al mismo tiempo la obtención de una acotación superior para la dimensión Hausdorff de los conjuntos generados por las series. Sea Σa_i una serie que cumple las condiciones i) e ii) del anterior punto. Sea s la dimensión de Hausdorff del conjunto C_a . Entonces se verifica:

$$i) s \leq \liminf_{i \rightarrow \infty} \frac{\log \rho_i}{\log R_i} = K$$

Para demostrarlo, consideremos un t cualquiera mayor que K . Tomemos un r tal que $K < r < t$. Fijemos un δ . Tomando i suficientemente avanzado, el recubrimiento de C_a por copias de C_a , resultará ser

un δ -recubrimiento. Para este recubrimiento tendremos:

$$\sum_{i \in P_i} \left\{ \sum_{n \in I} a_n + C_a \right\} \cap I^\Gamma \leq \rho_i(2R_i)^\Gamma = 2^\Gamma \rho_i(R_i)^\Gamma$$

Tomando i suficientemente avanzado, es posible conseguir:

$$-\frac{\log \rho_i}{\log R_i} < r$$

Por lo que si llamamos s_i al primer miembro de esa desigualdad, y teniendo en cuenta que para i suficientemente avanzado, $R_i < 1$:

$$\sum_{i \in P_i} \left\{ \sum_{n \in I} a_n + C_a \right\} 1^r < 2^r \rho_i (R_i)^{s_i} = 2^r$$

Esto prueba que $H^r_\delta(C_a) \leq 2^r$, y haciendo $\delta \rightarrow 0$, $H^r(C_a) \leq 2^r$. Esto significa que $s < t$, ya que en caso contrario, siendo $r < t$, debería ser $H^r(C_a) = \infty$. Siendo t un número arbitrario mayor que K , debe ser $s \leq K$. ■

5.3. EJEMPLOS

5.3.1. Series geométricas reales

El ejemplo más elemental de serie fractal es la serie geométrica real. Con este tipo de serie se obtiene fácilmente el conjunto de Cantor. Más concretamente, el conjunto de Cantor es el conjunto C_a asociado a la serie $\sum a_n$ donde $a_n = (2/3)(1/3)^n$. Para verlo, basta observar que cada posible suma de términos de esta serie puede ser interpretada como la expresión decimal en base tres de un número tal que todas sus cifras son ceros (si el correspondiente término de la serie no figura en la suma) o doses (si el correspondiente término de la serie figura en la suma). Si en lugar de tomar la razón igual a $1/3$ se toman otras razones (a continuación veremos que deben ser menores que $1/2$), se obtienen otros conjuntos llamados de tipo Cantor.

Vamos a probar el siguiente teorema:

i) La condición necesaria y suficiente para que una serie geométrica de razón r sea fractal, es que $0 < |r| < 1/2$

Sea la serie $\sum a_i$ donde $a_i = r^i$. La condición es suficiente, puesto que la condición i) de 5.2.5 resulta satisfecha, mientras que para esta serie es:

$$R_i = \frac{1 - r^{i+1}}{1 - |r|}$$

Por tanto si $|r| < 1/2$, $2^i R_i \rightarrow 0$.

La condición es necesaria, ya que $|r| > 1/2$ implica $L_1(C_a) = 1/(1-|r|)$. Para probarlo tratamos separadamente el caso $r > 0$ y el caso $r < 0$. Supongamos primero que $r > 0$. Observemos que como $r = 1/2 + \alpha$, $\alpha > 0$ se tiene:

$$\sum_{n \geq i} a_n = \frac{(1/2 + \alpha)^{i+1}}{1/2 - \alpha} > (1/2 + \alpha)^i = a_i$$

Es decir, la suma de todos los términos de la serie posteriores a un término dado, es mayor que este término. Es fácil entonces obtener una suma de términos de la serie igual a cualquier número comprendido entre cero y la suma total de la misma: Sea β tal número. Sea entonces i_1 el menor natural tal que:

$$\sum_{n=0}^{i_1} a_n > \beta$$

Sea $\beta_1 = \sum_{n=0}^{i_1-1} a_n$. Se sabe que $\beta - \beta_1 < a_{i_1}$ por lo que $\beta - \beta_1$ es superado por la suma de todos los términos posteriores al a_{i_1} .

Buscamos el menor natural i_2 que verifique:

$$\sum_{n=i_1+1}^{i_2} a_n > \beta - \beta_1$$

Y llamamos β_2 a la suma de los términos del i_1+1 al i_2-1 .

Procediendo de esta manera sucesivamente, es claro que la suma de todos los términos de la serie excluyendo el i_1 , i_2 , etc, es precisamente igual a β .

Si $r < 0$, entonces C_a es un intervalo $[p, q]$ con $p = r/(1-r^2)$ y $q = 1/(1-r^2)$. Obsérvese que p es la suma de todos los términos negativos de la serie y q la suma de todos los términos positivos. La longitud total del intervalo es $1/(1+r) = 1/(1-|r|)$, como en el caso anterior. Tomemos $\beta \in [p, q]$. Supongamos que β es positivo. Tomamos la suma de los primeros términos positivos de la serie, hasta el primero con cuya suma se rebase β . Se tendrá $\beta \in [p_1, q_1]$, donde p_1 es suma de términos positivos de la serie y $q_1 - p_1$ es un término de la serie, digamos, el a_{i_1} .

Razonando como en el caso anterior, la suma de los valores absolutos de todos los elementos posteriores al a_{i_1} rebasa a este en

valor absoluto, de forma que o bien p_1 mas la suma de todos los términos positivos de la serie posteriores al a_{i_1} rebasa a β , o bien

q_1 mas la suma de todos los términos negativos posteriores al a_{i_1}

es inferior a β . En ambos casos volvemos a conseguir situar a β en un intervalo $[p_2, q_2]$ que tiene por longitud un término de la serie, y por extremos una suma finita de términos de la serie. Prosiguiendo este proceso, se puede obtener β como suma de términos de la

serie. ■

5.3.2. Series geométricas complejas

Sea ahora $\sum a_i$ con $a_i = r^i$, $r \in \mathbb{C}$. Entonces se obtiene:

i) Una condición suficiente para que $\sum a_i$ sea fractal, es que $|r| < \sqrt{2}/2$.

Como $R_i^2 = (|r|^{i+1}/(1-|r|)^2) = (|r|^2)^{i+1}/(1-|r|)^2$, es fácil ver que si $|r| < 1/2$, $2^i \cdot R_i^2 \rightarrow 0$.

Naturalmente que una serie real de razón comprendida entre $1/2$ y $\sqrt{2}/2$, siempre puede ser considerada como compleja, satisfaciendo además la condición suficiente, y sin embargo sabemos que para esta serie, C_a es un intervalo. Para evitar este tipo de casos degenerados, es para lo que conviene, como señalamos en 5.2.4, considerar que las series con las que trabajamos están incluidas en el espacio \mathbb{R}^n de dimensión mínima posible.

También conviene destacar que en este caso no se puede afirmar que la condición suficiente señalada sea necesaria, quedando abierto aquí el interesante problema de determinar si el conjunto C_a tiene o no medida Lebesgue positiva cuando $|r| > \sqrt{2}/2$.

5.3.3. Autosemejanza para series geométricas

Sea $\sum a_i$ una serie geométrica, que supondremos real o compleja, y sea r su razón y a_0 su primer término. Sabemos que:

$$C_a = \{ a_0 + C_a \} \cup C_a$$

Por otra parte $C_{a_1} = r \cdot C_a$, por lo que la igualdad de arriba se escribe:

$$C_a = \{ a_0 + r C_a \} \cup r C_a$$

Consideremos ahora las semejanzas:

$$\Psi_1(x) = a_0 + rx ; \Psi_2 = rx$$

Tendremos:

$$C_a = \Psi_1(C_a) \cup \Psi_2(C_a)$$

Es decir, C_a es el conjunto de puntos invariante para las semejanzas Ψ_1 y Ψ_2 , ambas de razón r , por lo que basta que $|r| < 1$ para que sea satisfecha la condición a) de conjuntos autosemejantes dada en 5.1.3. Antes de completar la prueba de que, bajo ciertas restricciones para r , C_a es autosemejante, haremos un par de observaciones acerca de propiedades geométricas de C_a . Consideremos en primer lugar la biyección $\sigma : C_a \rightarrow C_a$ obtenida de la siguiente manera:

Si $x \in C_a$ es $x = \sum_{i \in I} a_i$, entonces $\sigma(x) = \sum_{i \in \mathbb{N}^* - I} a_i$

Se puede obtener también $\sigma(x)$ restando x a la suma total de la serie, es decir, $\sigma(x) = S - x$, donde $S = a_0/(1-r)$. El punto medio del segmento que une x con $\sigma(x)$ vale siempre $(\sigma(x) + x)/2 = S/2$, lo que prueba que σ es una simetría central con centro en el punto $S/2$. Así **hemos probado que los conjuntos C_a tienen una simetría central.**

También es interesante observar Ψ_1 es una semejanza de centro en el origen, consistente en el producto de un giro de ángulo $\arg(r)$ y centro en el origen compuesto con una homotecia de razón $|r|$ y el mismo centro, mientras que Ψ_2 tiene su centro en S , ya que $rS + a_0 = ra_0/(1-r) + a_0 = a_0/(1-r) = S$. Ψ_2 se obtiene componiendo un giro de ángulo $\arg(r)$ y centro en S con una homotecia de centro en S y razón $|r|$. En efecto, una transformación como la descrita se obtiene trasladando $-S$, multiplicando por r y trasladando S , es decir transforma x en $(x - S)r + S = xr + S(1-r) = xr + a_0 = \Psi_2$.

Inversamente, si tomamos dos puntos arbitrarios del plano y en cada uno de ellos una semejanza consistente en un giro por una homotecia centradas en el punto, de forma que los dos giros son del mismo ángulo α y las homotecias de la misma razón r , el conjunto invariante para estas dos transformaciones es, salvo traslación, el conjunto C_a asociado a una serie geométrica cuya razón es el número complejo $r = r_\alpha$ y cuya suma total es la diferencia de los dos puntos.

Probaremos ahora:

i) Una condición suficiente para que el conjunto C_a asociado a una serie geométrica compleja de razón r sea autosemejante, es que $0 < |r| < 1/2$.

Para probarlo, demostraremos que se cumple la condición de abierto (ver ii de 5.1.4). Para ello, es suficiente probar que son disjuntos los conjuntos $\Psi_1(C_a)$ y $\Psi_2(C_a)$, ya que entonces, tomando el ϵ -cuerpo paralelo (abierto) a C_a , es decir:

$$V_\epsilon = \{x \in C: d(x, C_a) < \epsilon\}$$

Para un ϵ suficientemente pequeño, V_ϵ verifica la condición de abierto.

El conjunto $\Psi_1(C_a)$, está formado por todas las sumas de términos de la serie que contienen a a_0 , mientras que $\Psi_2(C_a)$ está formado por las sumas de términos que no contienen a_0 . La suma de las normas de todos estos últimos términos es $|a_0| |r|/(1 - |r|)$, que, teniendo en cuenta la elección de r , es menor que $|a_0|$. Esto prueba que la diferencia entre una suma de términos que contiene a a_0 , y otra que no lo contiene, nunca se puede anular. ■

Aplicando además las consecuencias de la condición de abierto sobre

la dimensión del correspondiente conjunto autosemejante, se obtiene:

$$ii) \text{ Si } 0 < |r| < 1/2, \dim(C_a) = \log(2)/\log(1/|r|)$$

Basta ver que $\dim(C_a)$ es, según ii) de 5.1.3, aquel número s que verifica $|r|^s + |r|^s = 1$.

Con la condición impuesta a r , sólo pueden obtenerse fractales de dimensión menor que uno. En realidad, la condición $|r| < 1/2$, es demasiado restrictiva, incluso para conseguir que $\Psi_1(C_a) \cap \Psi_2(C_a)$ sea vacío, lo que tampoco es necesario para que haya autosemejanza. En los ejemplos del anexo puede observarse como para valores de r de norma mayor a $1/2$ no hay solapamiento (ver pgs. 144 y 145 para casos sin solapamiento. En la pag. 143, hay un ligero solapamiento), y en realidad parece probable que haya autosemejanza siempre que $|r| < \sqrt{2}/2$.

5.3.4. Otros ejemplos

Las series fractales son adecuadas para representar ciertos conjuntos de números con expresiones decimales especiales. Por ejemplo, la serie fractal $1/10 + 1/10 + 1/10 + 1/100 + 1/100 + 1/100 + 1/1000 + \dots$ da lugar al conjunto de todos los números en cuya expresión decimal en base 10 consta de cifras iguales a 0,1,2,3. Estudiaremos los fractales asociados a este tipo de series en un caso más general

Sean $a_1, a_2, a_3, \dots, a_k$ números naturales cuya suma vale s , y sea n un natural mayor que s . Consideremos la serie $a_1/n + a_2/n + \dots + a_k/n + a_1/n^2 + a_2/n^2 + \dots + a_k/n^2 + a_1/n^3 + \dots + \dots$. ¿Cuál es el fractal C_a asociado a esta serie. Se trata desde luego del conjunto de todos los números cuyas expresiones decimales en base n constan de las cifras que pueden obtenerse como sumas de números escogidos entre los a_i . Es fácil ver que el conjunto C_a es autosemejante, porque si todas las mencionadas sumas son b_1, b_2, \dots, b_m , la colección de semejanzas $\{\psi_i\}$, $0 \leq i \leq m$, tal que $\psi_i(x) = b_i/n + x/n$, verifica:

$$\bigcup_{i=1}^m \psi_i(C_a) = C_a$$

Además si $x \in \psi_i(C_a)$, tiene el término b_i/n , y si $y \in \psi_j(C_a)$, tiene el término b_j/n . Como la diferencia de estos términos vale al menos $1/n$, $x - y \neq 0$, ya que la suma de todos los términos de la serie $a_1/n^2 + a_2/n^2 + \dots$, vale exactamente $s/(n^2-1)$, y como $s < n-1$, resulta que dicha suma es menor o igual a $1/(n+1) < 1/n$, lo que significa que no puede ser compensada con estos términos la diferencia entre b_i/n y b_j/n . Esto prueba se cumple la condición de abierto según se ha razonado en i), y por tanto C_a es autosemejante. Se obtiene la siguiente curiosa relación entre la dimensión de Hausdorff de C_a y el número m de sumas distintas que pueden ser obtenidas con los a_i :

$$i) \dim(C_a) = \log m / \log(n)$$

5.4 FAMILIA DE SERIES FRACTALES ASOCIADA A UNA FUNCION ANALITICA

5.4.1.

Sea f una función analítica real o compleja en un abierto D . Suponemos además que f no es polinómica. Sea $x_0 \in D$ y consideremos el desarrollo en serie de f en x_0 :

$$f(x) = \sum_{i=0}^{\infty} a_i (x - x_0)^i$$

Sea ρ el radio de convergencia de la serie entera $\sum a_i X^i$. Afirmamos que:

i) Para todo $t \neq x_0$ perteneciente a una bola $B_r(x_0)$ centrada en x_0 y de radio $r = \rho/2$, la serie $\sum a_i (t-x_0)^i$ es una serie fractal.

La condición i) de 5.2.5 se cumple, porque siendo $t \neq x_0$, para que $a_i (t-x_0)^i$ se anule, se requiere que a_i lo haga. Pero como f no es un polinomio por hipótesis, infinitos valores de a_i deben ser no nulos.

La condición iii) también se cumple, ya que tomando $r_0 \in \mathbb{R}$ tal que $|t-x_0| < r_0 < r$ tendremos:

$$R_i = \sum_{n>i} |a_n| |t-x_0|^n < \sum_{n>i} |a_n| (r_0)^n$$

Multiplicando por 2^i se obtiene:

$$2^i R_i < 2^i \sum_{n>i} |a_n| (r_0)^n < \sum_{n>i} |a_n| (2r_0)^n$$

Como debido a la elección de r_0 es $2r_0 < r$ y r es el radio de convergencia de la serie entera $\sum a_n X^n$, esta debe ser convergente

en $X = 2r_0$, por lo que su resto, $\sum_{n>i} |a_n| (2r_0)^n$, tiende a cero. Esto

prueba que $2^i R_i \rightarrow 0$. ■

Puede conseguirse una mejora en la acotación para el radio de la bola para la cual las series $\sum a_i (t-x_0)^i$ son fractales, en el caso en que f sea compleja, si además suponemos que el radio de convergencia de $\sum a_i X^i$ sea $\rho \neq 1$. Las demás hipótesis son iguales que en el caso anterior, pero ahora se obtiene:

ii) Si $t \in B_r(x_0)$ y $t \neq x_0$, donde $r = (\rho/2)^{1/2}$

Observemos que como es $\rho \leq 1$, $\rho/2$, que era la acotación obtenida para r en el anterior desarrollo, es peor que $(\rho/2)^{1/2}$.

Para probarlo, basta demostrar, en lugar de $2^i R_i \rightarrow 0$, $2^i (R_i)^2 \rightarrow 0$ lo que resulta suficiente ahora que $\sum a_i \in C$, según 5.2.5. Tomamos como antes r_0 verificando $|t - x_0| < r_0 < r$. Tendremos:

$$\begin{aligned} R_i^2 &= (\sum_{n>i} |a_n| |t - x_0|^n)^2 < (\sum_{n>i} |a_n| (r_0)^n)^2 = \\ &= \sum_{m=0}^{\infty} (\sum_{p=0}^m |a_{i+1+p}| |a_{i+1+m-p}|) r_0^{2(i+1)+m} \end{aligned}$$

Si escribimos ahora $b_p = |a_{i+1+p}|$, tendremos:

$$R_i^2 < [\sum_{m=0}^{\infty} (\sum_{p=0}^m b_p \cdot b_{p-m}) r_0^m] r_0^{2(i+1)}$$

Ahora bien, la serie $\sum b_p (r_0)^p$ es convergente, y también es convergente la serie $\sum c_m (r_0)^m$ si es:

$$c_m = \sum_{p=0}^m b_p \cdot b_{p-m}$$

Ya que entonces:

$$\sum c_m (r_0)^m = (\sum b_p (r_0)^p)^2$$

Volviendo a las desigualdades anteriormente obtenidas, las escribimos ahora:

$$R_i^2 < (\sum c_m (r_0)^m) \cdot r_0^{2(i+1)}$$

Multiplícando por 2^{i+1} :

$$2^{i+1} R_i^2 < (2 \cdot r_0^2)^{i+1} \cdot (\sum c_m \cdot r_0^m) = (2 \cdot r_0^2)^{i+1} \cdot S$$

Donde S es la suma de la serie, que es convergente según se razonó más arriba. De $r_0 < (\rho/2)^{1/2}$ se deduce que $2 \cdot r_0^2 < \rho \leq 1$, por lo que $(2 \cdot r_0^2)^{i+1} \rightarrow 0$, y por tanto $2^i R_i^2 \rightarrow 0$. ■

5.3.2. Ejemplos

En el anexo de gráficos de este capítulo, se encuentran incluidos ejemplos de fractales obtenidos a partir de diferentes funciones analíticas.

La primera familia de fractales, correspondientes a distintas se-

ries geométricas complejas de razones menores a la unidad, pueden ser también entendidos como fractales generados por la función analítica compleja $f(z) = 1/(1-z)$, cuyo desarrollo en serie en el origen es $f(z) = \sum z^n$. El radio de convergencia de esta serie es 1, por lo que en si $t \in B_r(0)$, siendo $r = (1/2)^{1/2} = \sqrt{2}/2$, $\sum t^n$ es una serie fractal, resultado que ya conocíamos por el estudio hecho en 5.2.1 para series geométricas. En este ejemplo se observa la variedad de fractales que pueden obtenerse con una única función analítica, ya que según se sabe por dicho estudio, la dimensión del fractal asociado a una serie geométrica, depende del módulo de la razón.

En el mismo anexo figuran familias de fractales asociados a la función derivada de la anterior, $f'(z) = \sum n z^{n-1}$. La siguiente corresponde a la función $g(z) = \sum z^{n+1}/(n+1)$. En estas tres familias se puede observar como, a iguales valores de t se obtiene mayor masa en la función en cuyo desarrollo en serie el coeficiente correspondiente a igual potencia de z es mayor. La familia de fractales correspondiente a la función exponencial y en general a funciones trascendentes, son de estructura excesivamente fina para ser visualizados gráficamente. La siguiente familia de fractales representada, corresponde a series binómicas asociadas a la función $f(z) = (1+z)^\alpha$ para valores negativos de α .

5.3.3.

Procedemos ahora a examinar la dimensión de los fractales obtenidos por mediación de la función analítica real $f(x) = 1/(1-x)^2$ en un entorno del origen.

Sea $a_i(x) = (i+1)x^i$ el término general del desarrollo en serie en el origen de $f(x)$. Sabemos que si $0 < |t| < 1/2$, la serie de término general $a_i(t) = (i+1)t^i$ es una serie fractal, y llamaremos $C_a(t)$ al conjunto fractal asociado a esta serie. Nos proponemos probar el siguiente teorema:

i) Si $0 < |t| < 1 - (1/2)^{1/3}$, y s es la dimensión de Hausdorff de $C_a(t)$, entonces:

$$s = \frac{\log 2}{\log(1/|t|)}$$

La acotación superior se obtiene utilizando el resultado 5.2.7. La acotación inferior se obtiene construyendo una aplicación lipschitziana de $C_a(t)$ sobre el fractal autosemejante generado por la serie geométrica de razón t . Procedamos a probar el resultado, dividiendo la demostración en sucesivas etapas. La primera será:

a) Si $0 < |t| < 1 - \sqrt{2}/2$ entonces $\forall i \in \mathbb{N}^*$:

$$|a_i(t)| > R_i(t) = \sum_{n>i} |a_n(t)|$$

Comencemos por probarlo cuando $i = 0$. Se trata de demostrar

$$1 > 2!t^1 + 3!t^2 + 4!t^3 + \dots$$

Basta probar que si $0 < t < 1 - \sqrt{2}/2$, entonces

$$[1] \quad 1 > 2t + 3t^2 + 4t^3 + \dots$$

El segundo miembro de esta desigualdad vale $1/(1-t)^2 - 1$, por lo que la desigualdad es equivalente a :

$$2 > 1/(1-t)^2$$

Desigualdad que es cierta si $0 < t < 1 - \sqrt{2}/2$

Para el caso general, basta demostrar que si $0 < t < 1 - \sqrt{2}/2$:

$$(i+1)t^i > (i+2)t^{i+1} + (i+3)t^{i+2} + (i+4)t^{i+3}$$

Dividiendo por el primer miembro se obtiene que esta desigualdad es equivalente a:

$$1 > \frac{i+2}{i+1} t + \frac{i+3}{i+1} t^2 + \frac{i+4}{i+1} t^3$$

Pero para todo $i, k \in \mathbb{N}^*$, se verifica $(i+k)/(i+1) \leq k$, por lo que esta desigualdad se desprende de la [1]. ■

A continuación vamos a probar un resultado análogo para la serie obtenida por derivación término a término de $\sum a_i(x)$.

b) si $0 < |t| < 1 - (1/2)^{1/3}$ entonces $\forall i \in \mathbb{N}^*$:

$$(i+2)(i+1)|t|^i > (i+3)(i+2)|t|^{i+1} + (i+4)(i+3)|t|^{i+2} + \dots$$

Se comienza probándolo como antes para $i = 0$. Basta probar que si $0 < t < 1 - (1/2)^{1/3}$:

$$[2] \quad 2 > 3.2 t + 4.3 t^2 + 5.4 t^3 + \dots$$

El segundo miembro de esta desigualdad vale $2/(1-t)^3 - 2$, por lo que la desigualdad es equivalente a :

$$2 > \frac{2}{(1-t)^3} - 2$$

Y esta se verifica siempre que $0 < t < 1 - (1/2)^{1/3}$

Para probar el caso general, basta probar que si $0 < t < 1 - (1/2)^{1/3}$:

$$(i+2)(i+1)t^i > (i+3)(i+2)t^{i+1} + (i+4)(i+3)t^{i+2} + \dots$$

Que dividiendo por el primer miembro, es equivalente a:

$$[3] \quad 1 > \frac{(i+3)(i+2)}{(i+2)(i+1)} t + \frac{(i+4)(i+3)}{(i+2)(i+1)} t^2 + \frac{(i+5)(i+4)}{(i+2)(i+1)} t^3 + \dots$$

Pero teniendo en cuenta que $\forall i \in \mathbb{N}^*, K \geq 3$

$$[4] \quad \frac{(i+K)(i+K-1)}{(i+2)(i+1)} \leq \frac{K(K-1)}{2}$$

Dividiendo ambos miembros de la desigualdad [2] entre 2, utilizando a continuación la [4], se obtiene la [3]. ■

El siguiente paso será

c) Si p_i es la función de repetición (ver 5.2.6) asociada a $\Sigma a_i(t)$, y es $0 < |t| < 1 - \sqrt{2}/2$, tenemos:

$$p_i = 2^{i+1} \quad \forall i \in \mathbb{N}^*$$

Se trata simplemente de probar que todas las sumas de términos de la serie formadas con los i primeros términos, son siempre diferentes. La desigualdad probada en a) es la razón que nos permite asegurarlo. Sean P_i las partes del conjunto $\{0, 1, 2, \dots, i\}$, y sean $I, J \in P_i$, $I \neq J$. Entonces será $I \Delta J \neq \emptyset$. Consideremos el índice $k = \min\{I \Delta J\}$. Consideremos la expresión:

$$\Sigma_{I \in P_i} a_i(t) - \Sigma_{J \in P_i} a_i(t)$$

El primer término que no se anula es $a_k(t)$, pero según se ha probado en a), entre todos los términos de la serie que siguen a $a_k(t)$, sumados en valor absoluto, no llegan a sumar el valor absoluto de $a_k(t)$, por lo que la expresión no se puede anular. ■

Podemos ahora obtener la acotación superior para la dimensión:

d) si $0 < |t| < 1 - \sqrt{2}/2$ y s es la dimensión Hausdorff de $C_a(t)$:

$$s \leq \frac{\log 2}{\log(1/|t|)}$$

Para demostrarlo, basta utilizar el resultado de 5.2.7. Según el citado resultado, tendremos:

$$[5] \quad s \leq \liminf -\log p_i / \log R_i$$

Sabemos por c) que $p_i = 2^{i+1}$, y además, utilizando el hecho de que es $\forall i \in \mathbb{N}^* a_i(x)$ es la derivada de $b_{i+1}(x) = x^{i+1}$, y que la serie $\Sigma b_i(x)$ se puede derivar término a término si $t \in (-1, 1)$, poniendo:

$$S_{i+1}(t) = \sum_{n>i+1} b_n(t) = \frac{t^{i+2}}{1-t}$$

$$T_i(t) = \sum_{n>i} a_n(t)$$

Resulta :

$$T_i(t) = (S_{i+1}(t))' = \frac{t^{i+1}[(i+2)(1-t) + t]}{(1-t)^2}$$

Si como de costumbre llamamos $R_i(t)$ a la suma $\sum_{i \geq n} a_i(t)$, se tiene que $R_i(t) = T_i(|t|)$, por lo que sustituyendo en [5] p_i y $R_i(t)$ por los valores hallados, se obtiene:

$$\begin{aligned} s \leq \liminf - \frac{(i+1) \log 2}{(i+1) \log |t| + \log [(i+2)(1-|t|) + |t|] - \log(1-|t|) + \log 2} = \\ = \frac{\log 2}{\log(1/|t|)} \end{aligned}$$

Que era la desigualdad deseada. ■

Para concluir la demostración del teorema consideremos la función $g(x) = x/(1-x)$, analítica en un entorno del origen de radio 1. El término general de su desarrollo en serie es $b_i(x) = x^{i+1}$. Consideremos la serie de término general $b_i(t) = t^{i+1}$, que es una serie fractal si $0 < |t| < 1/2$, y tiene asociado un conjunto fractal autosemejante $C_b(t)$ cuya dimensión de Hausdorff es, según lo probado en ii) de 5.3.3 $\log 2 / \log(1/|t|)$. Probaremos el siguiente resultado:

e) Si $0 < |t| < 1 - (1/2)^{1/3}$, existe una aplicación lipschitziana y sobreyectiva $\Psi : C_a(t) \longrightarrow C_b(t)$.

La aplicación se define, si $I \in \mathbb{N}^*$:

$$\Psi \left(\sum_{i \in I} a_i(t) \right) = \sum_{i \in I} b_i(t)$$

Para asegurarnos de que se trata de una aplicación es preciso probar que todos los elementos de $C_a(t)$ tienen una representación única de la forma:

$$\sum_{i \in I} a_i(t), \quad I \in \mathbb{N}^*$$

Si $I, J \in \mathbb{N}^*$ son tales que $I \neq J$, entonces se prueba :

$$\left| \sum_{i \in I} a_i(t) - \sum_{i \in J} a_i(t) \right| \neq 0$$

Utilizando la propiedad a), y haciendo uso de la argumentación dada en iii) de 5.2.4.

Para demostrar que se trata de una aplicación lipschitziana, consideremos dos puntos $s_1, s_2 \in C_a(t)$. Sea para ciertos $I, J \subset N^k$:

$$s_1 = \sum_{i \in I} a_i(t) ; \quad s_2 = \sum_{j \in J} a_j(t)$$

Y consideremos las funciones:

$$f_1(x) = \sum_{i \in I} b_i(x) ; \quad f_2(x) = \sum_{j \in J} b_j(x)$$

Consideremos también la función $F(x) = f_1(x) - f_2(x)$. Aplicando el teorema del valor medio a esta función en el intervalo $[0, t]$, se obtiene:

$$F(t) = F'(x)t, \quad x \in (0, t)$$

Tomando valores absolutos:

$$|F(t)| = |F'(x)| |t|$$

Como las series que definen $f_1(x)$ y $f_2(x)$ se pueden derivar indefinidamente término a término en el intervalo $(-1, 1)$, resulta que:

$$F''(x) = \sum_{i \in I} b_i''(x) - \sum_{j \in J} b_j''(x), \quad x \in (-1, 1)$$

Ahora bien, si $I \neq J$, utilizando el resultado de b) y el razonamiento efectuado en iii) de 5.2.4, tal como se ha hecho más arriba, se obtiene que el segundo miembro de la igualdad es distinto de cero si $0 < |t| < 1 - (1/2)^{1/3}$, de donde si t satisface estas condiciones:

$$|F'(x)| \leq |F'(t)|$$

Lo que significa:

$$|F(t)| \leq |F'(t)| |t|$$

La prueba concluye sin más que observar que $F(t) = \Psi(s_1) - \Psi(s_2)$, y que $F'(t) = s_1 - s_2$, con lo que se obtiene:

$$|\Psi(s_1) - \Psi(s_2)| \leq |s_1 - s_2| (1 - (1/2)^{1/3})$$

Que prueba que Ψ es lipschitziana. Que Ψ es sobreyectiva, es evidente. ■

Finalmente probaremos la acotación inferior de la dimensión:

f) Si $0 < |t| < (1 - (1/2)^{1/3})$ y s es la dimensión Hausdorff de $C_a(t)$:

$$\log 2 / \log(1/|t|) \leq s$$

Esto es consecuencia de e) y de un conocido lema (ver ref. [F1]) que afirma que si existe una aplicación lipschitziana y sobreyectiva $\Psi : E \rightarrow F$, donde E y F son subconjuntos de un espacio métrico, entonces:

$$H^s(F) \leq H^s(E)c^s$$

donde c es la constante de Lipschitz de Q . Naturalmente esto entraña que $\dim(F) \leq \dim(E)$. En nuestro caso $F = C_b(t)$ tiene dimensión $\log 2 / \log(1/|t|)$ según se ha señalado ya, y de aquí se deduce el resultado. Queda concluida la prueba de f), y también del teorema. ■

Este resultado se puede generalizar a todas las funciones de la forma $1/(1-x)^K$, $K \in \mathbb{N}$, obteniéndose así el siguiente resultado:

ii) Si $k \in \mathbb{N}$, $0 < |t| < 1 - (1/2)^{1/(k+1)}$ y $\sum a_j(t)$ es una serie de término general $1/(1-t)^k$, el fractal $C_a(t)$ asociado a esta serie tiene dimensión $\log 2 / \log(1/|t|)$.

Para probarlo se realizaría inductivamente. Sabiendo la dimensión de los fractales generados por $1/(1-x)^2$, obtenida en el teorema anterior, se construye una aplicación lipschitziana y sobreyectiva de los generados por $1/(1-x)^3$ sobre los generados por $1/(1-x)^2$ de la forma indicada en el paso f) de la anterior demostración, y esto nos proporciona la acotación inferior. El paso b) ahora se consigue por idéntico proceso, pero imponiendo $0 < |t| < 1 - (1/2)^{1/4}$. La acotación superior se consigue como en d), y esto permite obtener la conclusión en el caso $k = 3$, procediéndose inductivamente con el resto de los valores de k .

5.5. OTROS ALGORITMOS PARA LA GENERACION DE FRACTALES .

En este apartado se muestra como pueden obtenerse fractales por procedimientos análogos a los descritos en el apartado 2, pero utilizando otros algoritmos.

5.5.1. Obtención de fractales mediante productos infinitos.

Sea el producto infinito:

$$\prod_{i=0}^{\infty} (1 + a_i)$$

Donde a_i es un número real o complejo, y supongamos que es absolutamente convergente.

Definimos :

$$C_a = \bigcup_{i \in \mathbb{N}^*} \prod_{i \in I} (1 + a_i)$$

Probaremos que si la serie $\sum a_i$ verifica las condiciones suficientes para serie fractal de 5.2.6, el producto infinito también es fractal, en el sentido de que se verifica que C_a tiene es un conjunto compacto y perfecto con potencia del continuo, y que $L_\alpha(C_a) = 0$, siendo $\alpha = 1, 2$, según que el producto sea real o complejo.

Sin pérdida de generalidad, podemos suponer que $1 + a_i \neq 0 \forall i \in \mathbb{N}^*$.

Probaremos primero:

i) Si para infinitos valores de i es $a_i \neq 0$, C_a tiene potencia del continuo, y es un conjunto compacto y perfecto.

Consideremos la serie $\sum \log(1 + a_i)$ donde \log representa la rama principal de la función logarítmica compleja. Dicha serie es absolutamente convergente y tiene infinitos términos no nulos, por lo que el conjunto C_b asociado a ella, tiene según iii) de 5.2.3 potencia del continuo. La aplicación biyectiva $\Psi : C_a \longrightarrow C_b$ definida por:

$$\Psi \left(\prod_{i \in I} (1 + a_i) \right) = \log \left(\prod_{i \in I} (1 + a_i) \right) = \sum_{i \in I} \log(1 + a_i)$$

donde \log es siempre la rama principal de la función logarítmica compleja, prueba que C_a tiene potencia del continuo. Como Ψ es un homeomorfismo, y como C_b es un conjunto compacto por 5.2.2 y perfecto según ii) de 5.2.4, obtenemos el resultado. ■

Sea ahora p_i la función de repetición asociada a la serie $\sum a_i$ y R_i , como usualmente, $\sum_{n>i} a_n$, y $\alpha = 1, 2$ según la serie sea real o compleja. Se cumple:

ii) Si $p_i \cdot R_i^\alpha \longrightarrow 0$, entonces $L_\alpha(C_a) = 0$

Para probarlo, escribamos la propiedad correspondiente a la propiedad de bipartición en fractales asociados a series (ver 5.2.3). Ahora podremos escribir para cualquier $j \in \mathbb{N}^*$:

$$C_a = \bigcup_{i \in P_i} \{ \prod_{j \in I} (1 + a_j) \cdot C_{a_{i+1}} \}$$

Donde P_i son las partes de $\{0, 1, 2, \dots, i\}$, y $C_{a_{i+1}}$ representa todos los productos que pueden formarse con términos de la serie a partir del $i+1$, inclusive. Tendremos:

$$[1] \quad L_\alpha(C_a) \leq p_i \cdot p \cdot L_\alpha(C_{a_{i+1}})$$

Donde $p = \prod_{i \in \mathbb{N}^*} (1 + |a_i|)$. Tengamos en cuenta para ello que:

$$L_\alpha \left\{ \left(\prod_{j \in I} (1 + a_j) \right) \cdot C_{a_i} \right\} \leq p \cdot L_\alpha \{ C_{a_i} \}$$

Y también que la función de repetición asociada a la sucesión $1+a_1, 1+a_2, 1+a_3, \dots$, es la misma que la asociada a a_1, a_2, a_3, \dots .

Analicemos ahora el producto:

$$\prod_{n=i+1}^{\infty} (1 + a_n) = 1 + \sum_{n>i} a_n + \sum_{n>m>i} a_n a_m + \sum_{n>m>p>i} a_n a_m a_p + \dots$$

Si $x \in C_{a_i}$, desarrollando el producto de los factores que lo componen, obtenemos algunos de los sumandos del segundo miembro, entre los cuales siempre estará el 1. Por eso se puede afirmar, si $R_i = \sum_{n>i} a_n$ si tomamos i suficientemente avanzado para que $R_i < 1$:

$$\begin{aligned} |x - 1| &\leq \left| \sum_{n>i} a_n \right| + \left| \sum_{n>m>i} a_n a_m \right| + \dots \leq \\ &\leq R_i + R_i^2 + R_i^3 + \dots = R_i / (1 - R_i) \end{aligned}$$

Esta desigualdad prueba que los elementos de $C_{a_{i+1}}$ pertenecen a una bola de centro 1 y radio $R_i / (1 - R_i)$, luego:

$$L_\alpha(C_{a_{i+1}}) \leq [2R_i / (1 - R_i)]^\alpha \cdot C_\alpha$$

Lo que en nuestra hipótesis de que $p_i R_i \alpha \rightarrow 0$, sustituyendo en [1], y haciendo tender i a infinito, proporciona $L(C) = 0$. ■

En el anexo de este capítulo, se incluyen algunos fractales generados mediante productos infinitos de la forma señalada en este punto.

5.5.2.

Incluimos a continuación otros algoritmos de generación de fractales basados en ideas similares a las de los ya vistos. Haremos solamente una descripción ya que no hemos desarrollado todavía el análisis detallado de los mismos.

El algoritmo de generación de fractales por series, puede también mirarse bajo el siguiente punto de vista: Tomamos una serie absolutamente convergente $\sum a_i$, $a_i \in \mathbb{R}^n$, y consideramos el conjunto \mathbb{A}_2 de todas las sucesiones infinitas de ceros y unos, de manera que si $\alpha \in \mathbb{A}_2$, $\forall i \in \mathbb{N}^*$, $\alpha_i = 0$ ó $\alpha_i = 1$. Entonces:

$$C_a = \bigcup_{\alpha \in \mathbb{A}_2} \sum_{i \in \mathbb{N}^*} \alpha_i a_i$$

Coincide claramente con el conjunto asociado a la serie $\sum a_i$, tal como fué definido en 5.2.1

Un conjunto notable que puede ser obtenido mediante una representación similar, es el conjunto de Besicovitch, que tiene propiedades geométricas muy interesantes descritas por M. de Guzman. (ver ref [M2])

Dicho conjunto se construye (ver fig 5.1) retirando de un cuadrado el complementario de los cuatro cuadrados sombreados, y luego procediendo a hacer la misma operación en cada uno de ellos, y repitiendo este proceso indefinidamente.

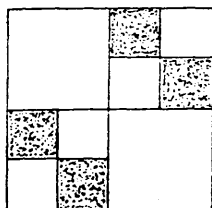


fig 5.1

No es difícil probar que dicho conjunto se representa mediante la expresión:

$$\bigcup_{\alpha \in \mathbb{Q}_2} \sum_{i=1}^{\infty} (1/2)^i \cdot (\alpha_i, \beta_i)$$

Donde $\beta_i = \alpha_i$ si i es impar, y $\beta_i = 1 - \alpha_i$, si i es par

Consideremos ahora el conjunto \mathbb{Q}_3 de todas las sucesiones formadas con los números 0, 1, -1, y la misma serie, y sea el conjunto:

$$C_a' = \bigcup_{\alpha \in \mathbb{Q}_3} \sum_{i \in \mathbb{N}^*} \alpha_i a_i$$

Es fácil ver que este conjunto coincide con el conjunto:

$$\{x - y : x, y \in C_a\}$$

Una condición suficiente para que su medida Lebesgue sea nula, será, según razonamientos similares a los desarrollados en 5.2.5, que $3^i (R_i)^n \rightarrow 0$. En el caso de que la serie sea geométrica de razón r , C_a' se escribirá:

$$C_a' = \{-a_0 + rC_a'\} \cup \{rC_a'\} \cup \{a_0 + rC_a'\}$$

Por lo que si ponemos $\Psi_1(x) = -a_0 + rx$, $\Psi_2(x) = rx$, $\Psi_3(x) = a_0 + rx$, se tendrá:

$$C_a' = \Psi_1(C_a) \cup \Psi_2(C_a) \cup \Psi_3(C_a)$$

Lo que prueba que C_a' es autosemejante si imponemos condiciones adecuadas a r . Es suficiente que $|r| < 1/2$ para evitar el solapamiento, (no será seguramente necesario), y en tal caso el fractal C_a' tendrá dimensión $\log 3 / \log(1/|r|)$.

El proceso descrito puede ampliarse en formas muy variadas. Por ejemplo, si sustituimos el conjunto \mathbb{Q}_3 por el \mathbb{Q}_5 , formado por sucesiones de los números -2, -1, 0, 1, 2, se obtendría el conjunto C_a'' , etc.

Más en general, podemos construir conjuntos "fractales" a través de la siguiente definición:

Sea como antes $\sum a_i$, $a_i \in \mathbb{R}^n \forall i \in \mathbb{N}^*$, una serie absolutamente convergente, y sea \mathbb{Q}_n el conjunto de sucesiones de los n elementos del conjunto $\langle 1, n \rangle = \{1, 2, 3, \dots, n\}$, y sea $\sigma : \langle 1, n \rangle \rightarrow K$ una aplicación inyectiva de $\langle 1, n \rangle$ en K . Suponemos que $K = \mathbb{R}$ en el caso general de que $a_i \in \mathbb{R}^n$, pero también puede ser $K = \mathbb{C}$ si $a_i \in \mathbb{C} \forall i \in \mathbb{N}^*$, y consideremos el conjunto:

$$C_a = \bigcup_{\alpha \in \mathbb{Q}_n} \sum_{i \in \mathbb{N}^*} \sigma(\alpha_i) a_i$$

De este conjunto cabe esperar propiedades fractales. En el caso de que la serie sea geométrica, imponiendo condiciones a r , estaríamos de nuevo frente a un conjunto autosemejante.

Otro ejemplo de fractal ya conocido de obtención por un método similar a los expuestos, es el descrito por Rogers (ver ref [R1]), a través del algoritmo de las fracciones continuas. Se trata del conjunto de fracciones continuas $\{a_0, a_1, a_2, \dots\}$, que pueden construirse de forma que $\forall i, a_i \in \{1, 2\}$. Rogers prueba que este conjunto es un conjunto de dimensión Hausdorff comprendida entre $1/3$ y $2/3$. Si convenimos en que para toda $\alpha \in \mathbb{R}_2$ es $\Psi(\alpha)$ la fracción continua $[a_0, a_1, a_2, \dots]$, entonces el conjunto F es simplemente $\Psi(\mathbb{R}_2)$.

En el anexo de gráficos a este capítulo, presentamos algunos conjuntos de apariencia fractal obtenidos mediante la siguiente definición

$$C_a = \bigcup_{i \in \mathbb{N}} \prod (z + a_i)$$

Siendo $\sum a_i$ una serie absolutamente convergente de números complejos, y siendo z un número complejo de módulo unidad y argumento racional. Téngase en cuenta que ahora el producto:

$$\prod_{i \in \mathbb{N}} (z + a_i)$$

Converge no a un punto, sino a varios puntos del plano si z tiene argumento racional, y a todo un círculo si el argumento fuera irracional. En nuestros ejemplos, hemos usado para $\sum a_i$ la serie geométrica, y para z complejos de argumento racional.

Cerramos esta memoria planteando la siguiente pregunta, a la que conducen de forma natural los anteriores ejemplos:

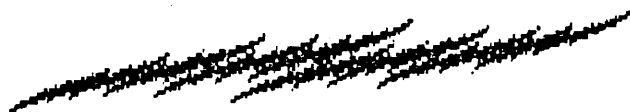
Sea una aplicación $\Psi: \mathbb{R}_n \rightarrow E$, siendo E un espacio métrico, ¿bajo que condiciones de Ψ puede esperarse que $\Psi(\mathbb{R}_n)$ tenga propiedades fractales? ¿Como influyen las propiedades de Ψ en las de $\Psi(\mathbb{R}_n)$?

Nota: Todos los ejemplos expuestos en este punto obedecen al esquema anterior, a excepción del último, en que la Ψ es una aplicación $\Psi: \mathbb{R}_n \rightarrow P(E)$

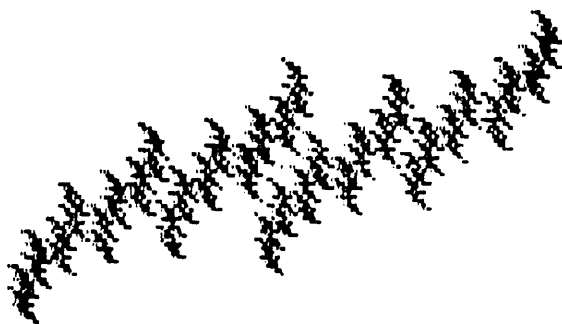


ANEXO DE GRAFICOS DEL

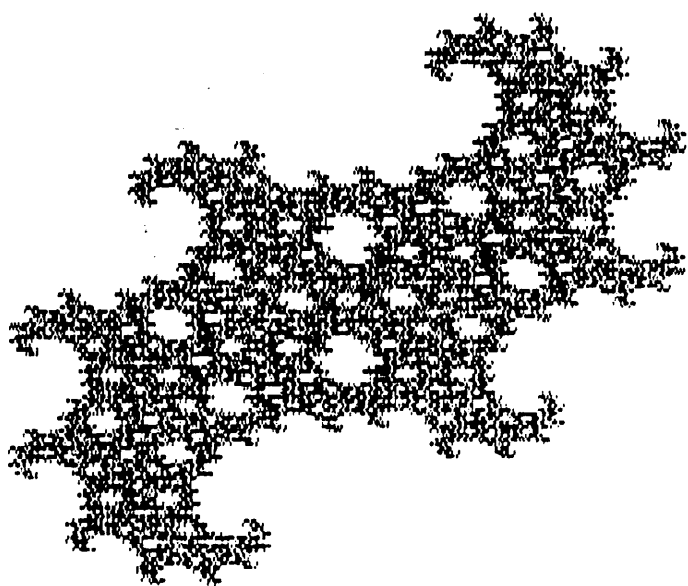
CAPITULO V



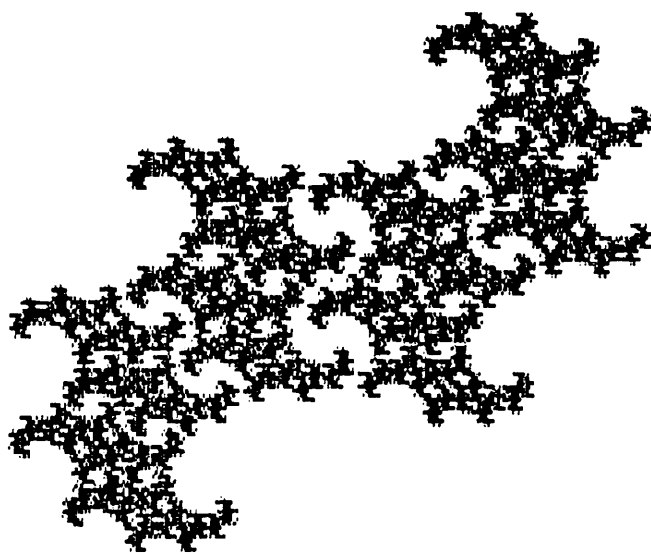
Serie : Geométrica
Razón : $.66_{10}$ (ángulo en grados)



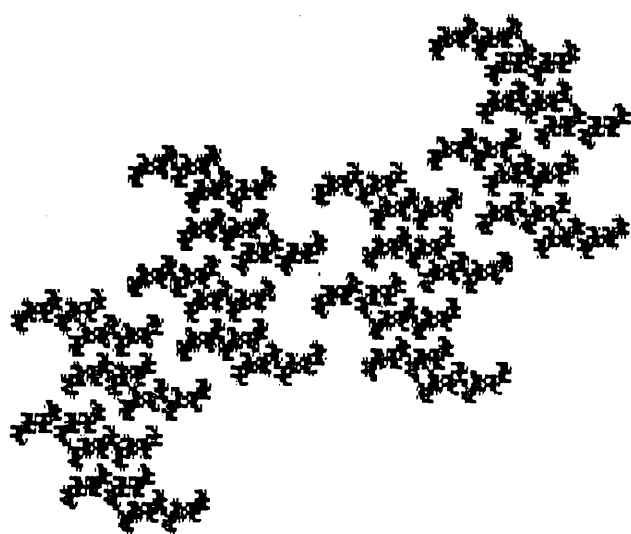
Serie : Geométrica
Razón : $.66_{15}$



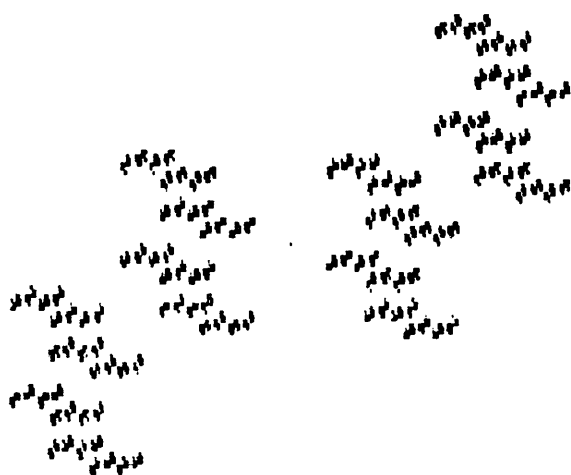
Serie: Geométrica
Razón: .730



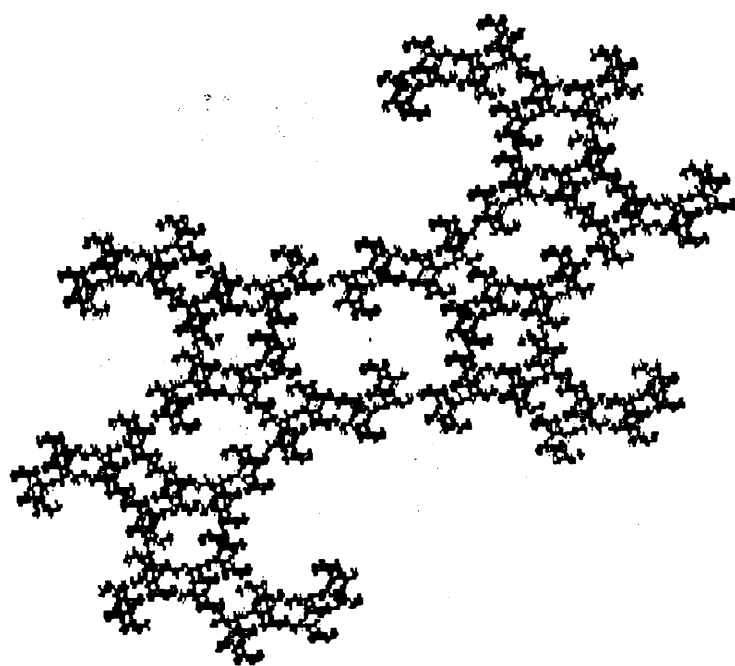
Serie: Geométrica
Razón: .6830



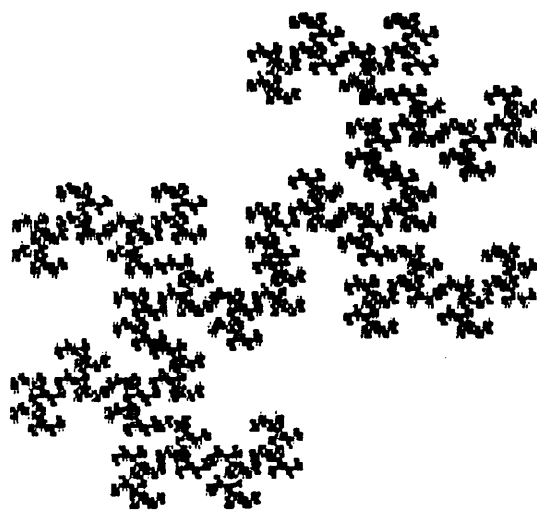
Serie : Geométrica
Razón : .6630



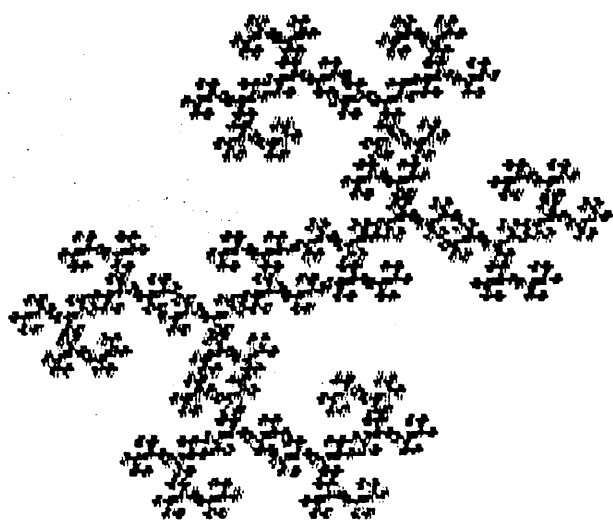
Serie: Geométrica
Razón: .6030



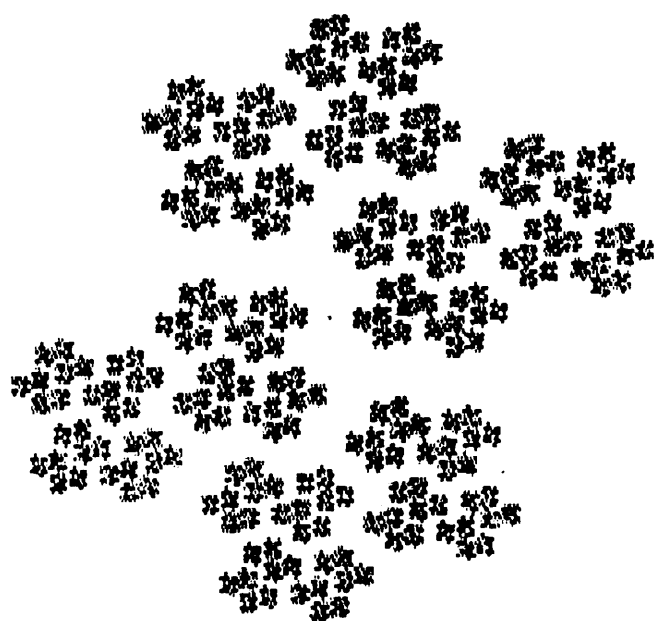
Serie: Geométrica
Razón: .6640



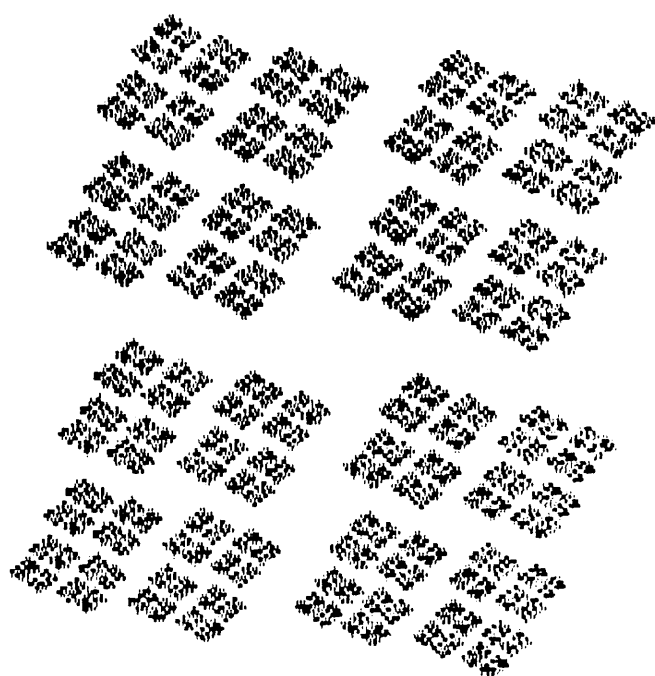
Serie: Geométrica
Razón: .6645



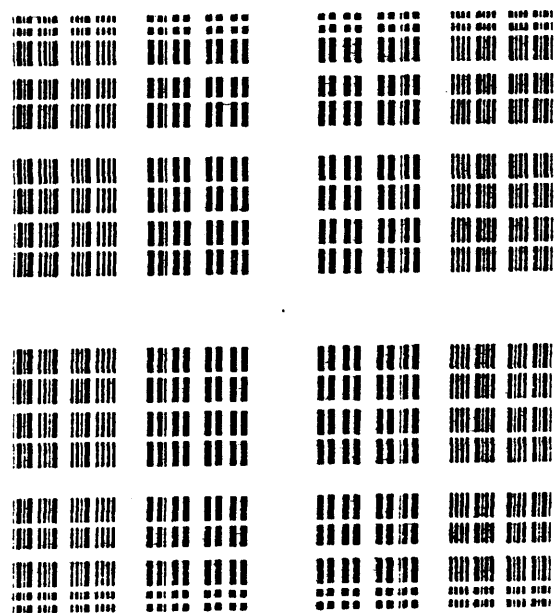
Serie: Geométrica
Razón: .6660



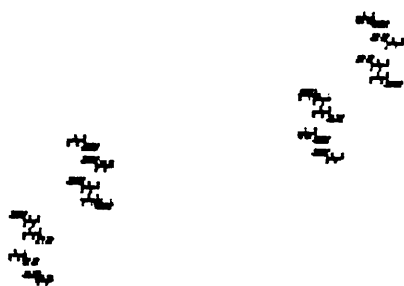
Serie: Geométrica
Razón: .6670



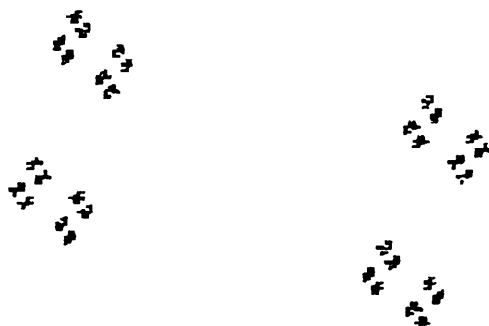
Serie: Geométrica
Razón: .6685



Serie: Geométrica
Razón: .6690



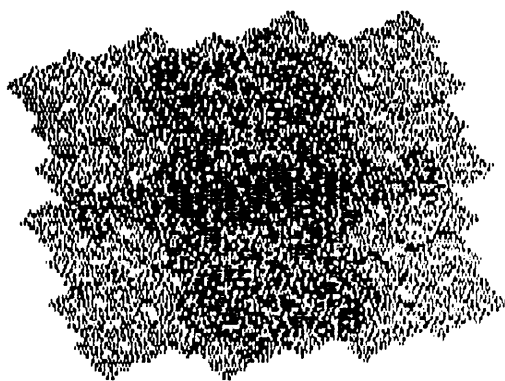
Serie: $\Sigma z^{n+1}/(n+1)$
 $z = .6630$



Serie: $\Sigma z^{n+1}/(n+1)$
 $z = .6680$

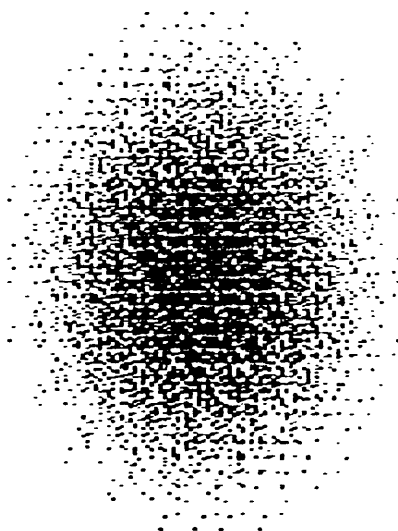
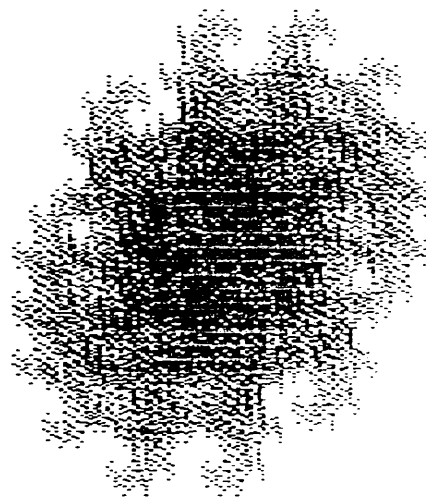


Función : $f(x) = (1 + x)^\alpha$
 $\alpha = -1.5$
 $x = .6630$

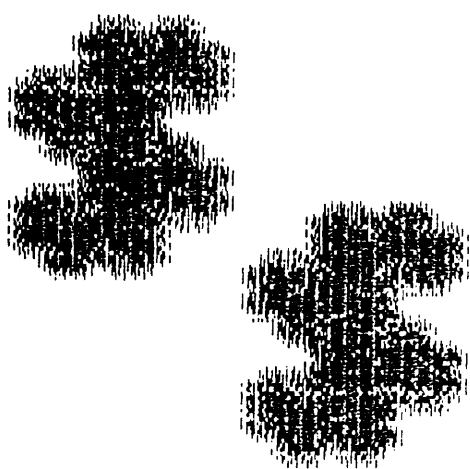


Funci3n: $f(x) = (1 + x)^\alpha$
 $\alpha = -1.5$
 $x = .6685$

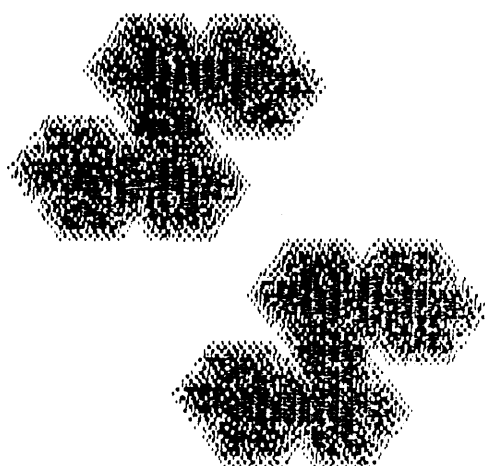
Función: $f(x) = (1 + x)^\alpha$
 $\alpha = -2$
 $x = .6630$



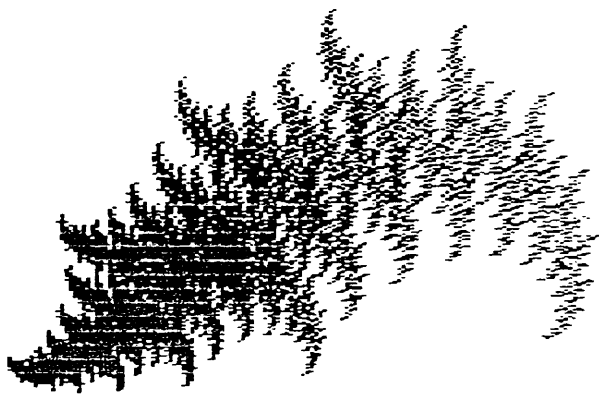
Función: $f(x) = (1 + x)^\alpha$
 $\alpha = -5$
 $x = .6630$



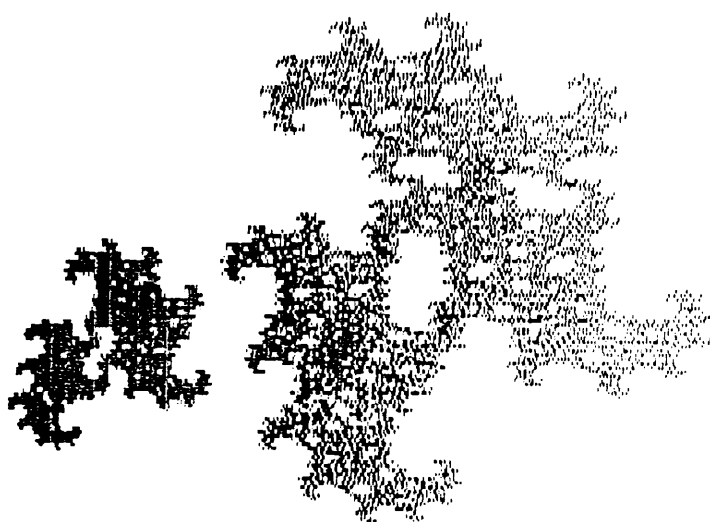
Función: $f(x) = \log(1 + x)$
 $x = .945$



Función: $f(x) = \log(1 + x)$
 $x = .960$



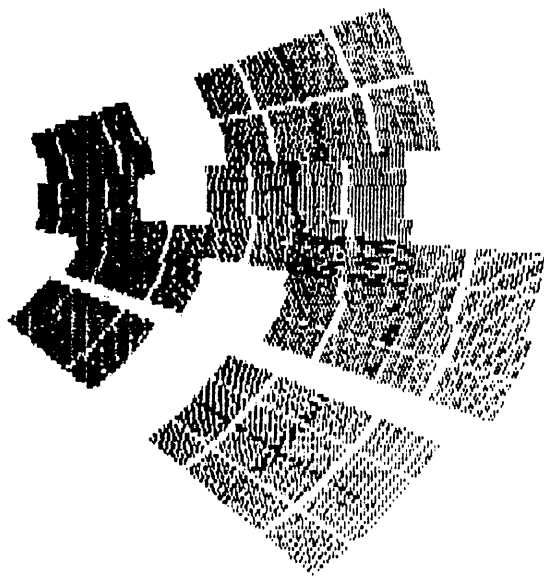
Producto infinito
Término general: $(1 + z^n)$
 $z = .715$



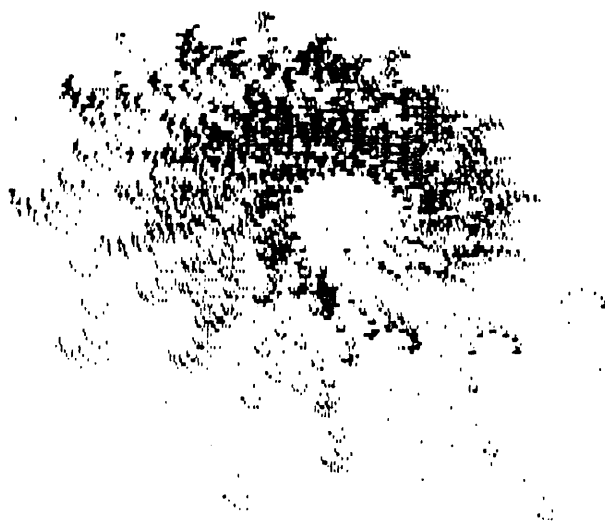
Producto infinito.
Término general: $(1 + z^n)$
 $z = .730$



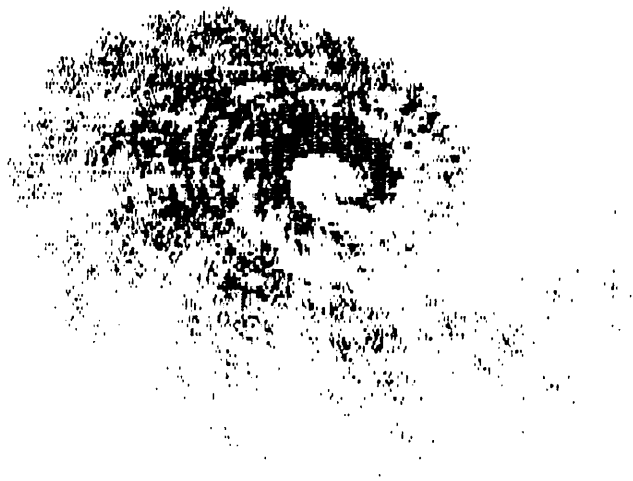
Producto infinito
 Término general: $(1 + z^n)$
 $z = .780$



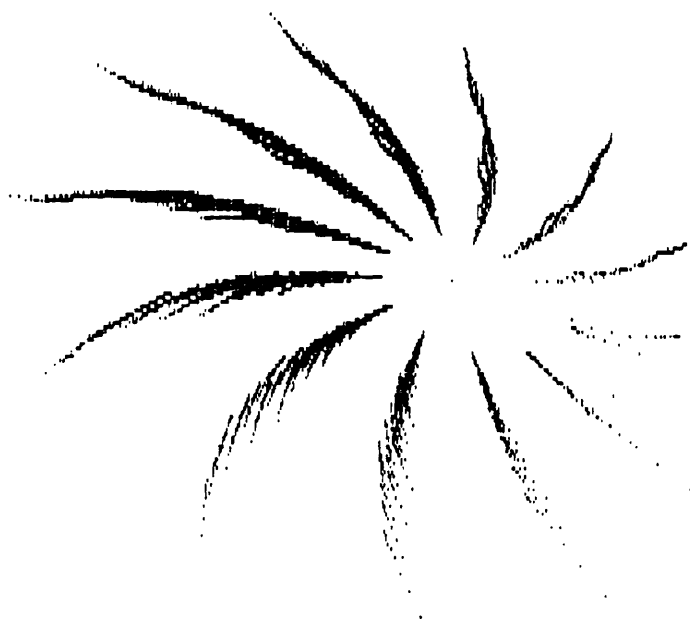
Producto infinito
 Término general: $(1 + z^n)$
 $z = .790$



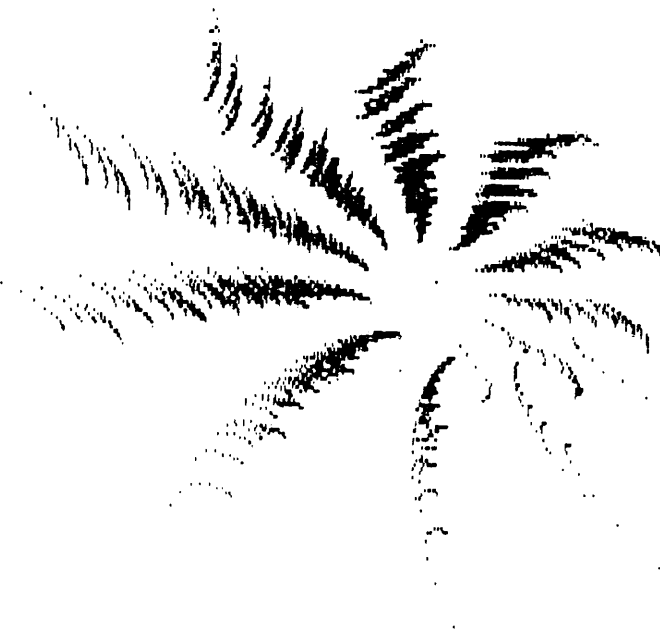
Producto infinito
Término general: $(u + z^n)$
 $u = 130$
 $z = .730$



Producto infinito
Término general: $(u + z^n)$
 $u = 130$
 $z = .790$



Producto infinito
Término general: $(u + z^n)$
 $u = 130$
 $z = .70$



Producto infinito
Término general $(u + z^n)$
 $u = 130$
 $z = .710$

BIBLIOGRAFIA

- [A1] V. I. Arnold. Problemes ergodiques de la m canique clasique
Gauthier-Villars editeur. Paris 1967
- [B1] G. Birkhoff. Dynamical Systems.
Am. Mat. Soc. Colloquium Publications, Vol IX. Providence 1927
- [F1] K. J. Falconer. The Geometry of fractal sets.
Cambridge University Press 1985.
- [G1] G. A. Gal'perin. On the existence of nonperiodic and nondense
billiard trajectories in polygons and polyedra.
Soviet. Math. Dokl. Vol 29 (1984). N  2.
- [G2] M. Guzm n. Real variable methods in Fourier Analysis.
North Holland Mathematics Studies. New York 1981.
- [G3] M. Guzm n. Differentiation of integrals in R^n .
Springer Lecture Notes in Mathematics.
- [G4] M. Guzm n y Baldomero Rubio. Integraci n: Teor a y t cnicas.
Editorial Alhambra. Madrid 1979.
- [H1] G. H. Hardy and E. M. Wright. An introduction to the Theory of
Numbers.
Oxford Univerity Press, Oxford 1968.
- [I1] M. C. Irwin. Smooth Dynamical Systems
Univerity of Liverpool. Academic Press 1980.
- [K1] M. Keane. Interval Exchange Transformatios.
Math. Z. 141. 25 - 31 (1975)
- [K2] M. Keane, C. Boldrighini and F. Marchetti. Billiards in
polygons.
The Annals of Probability, 1978, Vol 6 n  4, 532-540.
- [K3] Michael Keane. Non ergodic interval exchange transformations.
Israel Journal of Mathematics. Vol. 26 n  2 1977.
- [K4] B. Keynes and Newton. A "mininal" non uniquely ergodic
interval exchange transformation.
Math. Z. 148, 101-105 (1976)
- [K5] S. Kerckhoff, H. Masur and J. Smillie. Ergodicity of billiard
flows and Quadratic Differentials.
- [M1] R. Ma  . Ergodic Theory and Differentiable Dynamics.
Ergebnisse der Mathematik und ihrer Grenzgebiete.
3. Folge Band 8, Springer Verlag 1987.
- [M2] H. Masur. Interval Exchange Transformations and Measured
Foliations.
Annals of Mathematics 115 (1982) 169-200.

- [M3] B. Mandelbrot. Los objetos fractales.
Tusquets editores. Barcelona 1987.
- [M4] J. Marion. Mesure de Hausdorff d'un fractal a similitude interne.
Ann. Soc. Math. Quebec, 1986.
- [R1] C. A. Rogers. Hausdorff Measures.
Cambridge University Press, 1970
- [S1] I. J. Schoenberg. Mathematical Time Exposures.
Math. Assoc. of Am. Washington, 1982
- [S2] Ya. G. Sinai. Introduction to Ergodic Theory.
Mathematical Notes. Princeton University Press, 1977.
- [S3] L. A. Santaló. Geometría proyectiva.
E.U.D.E.B.A., Buenos Aires 1977.
- [S4] S. Sternberg. Celestial Mechanics.
W. A. Benjamin. Inc. 1989.
- [T1] G. Targonsky. Topics in Iteration Theory.
Vanderhoeck & Ruprecht in Gottingen, 1981.
- [V1] W. Veech. Interval Exchange Transformations.
Journal D'Analyse Mathématique. Vol. 33 (1978).
- [W1] Weyl.
Mathematical Annalen, 77 (1916), 312-52